Using student views to develop a formative assessment policy in a sixth-form History department

Emma Hawkins

A Research & Development Project Submitted for the MSc Learning & Teaching 2016

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	Hawkins
First Name	Emma
Faculty Board	Department of Education
Title of Dissertation	Using student views to develop a formative assessment policy in a sixth-form History department

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.

Using student views to develop a formative assessment policy in a sixth-form History department.

Using student views to develop a formative assessment policy in a sixth-form History department.

Abstract

This project sought to explore the views of sixth-form history students concerning the efficacy of their department's formative assessment policy, and builds on my part 2 research into student views of self and peer marking. The project first consisted of a literature review in order to provide the grounding needed to adapt the department's current assessment policy. This policy was then implemented by all staff in the department at the start of the academic year. Questionnaires and group interviews were carried out to explore how the students responded to this policy, with the intention that their responses would then be mediated by staff to produce an adapted policy for the start of the next academic year.

One of the key findings of this project was that despite increased student confidence in self-assessment, students still reported considerable reliance on the teacher. Among other conclusions drawn, it will ultimately be suggested that to further shift the balance of power in assessment away from the teacher, students need greater ownership not only over the assessment of their work, but also the assessment policy itself. Therefore, although the original aim of the project was for the staff to redraft the department's assessment policy based on student views, instead the outcome developed into a decision to give students a role in regularly revising the policy, encouraging a more organic and embedded approach.

Although this approach to policy-making in general is far from novel, it is rarely applied specifically to formative assessment, where educationalists tend to empower students within the narrow boundaries of teacher-determined assessment activities. It is still often the teacher ultimately who 'owns' and 'provides' formative assessment activities, with students required only to play their ascribed role. Therefore, this approach contrasts considerably with some of the contemporary literature that sees formative assessment as a concern of teacher pedagogy rather than student ownership.

Rationale

'There is a lot known about feedback, but there is much more to be discovered about how to optimize its power in the classroom' (Hattie, 2012: 134).

Research priorities

It appears that for many educationalists, the case has been made for the value of formative assessment, and the main task for researchers now lies in effectively applying it to the classroom context. The challenge is perhaps best summarized by Hattie: '*That feedback is critical to raising achievement is becoming well understood, but that it is so absent in classrooms (at least in terms of being received by students) should remain an important conundrum'* (2012: 135).

In this project, I have aimed to explore within my own classroom context how I can better embed formative assessment practices by exploring how students experienced formative assessment this year. Whilst my part 2 project explored student views of self and peer assessment, it went no further than exploring their opinions. My part 3 would extend this by using student views to bring the students into the policy-making process. The project focuses primarily on self-assessment, as this was highlighted by my part 2 research as being an area that students were more receptive to than peer assessment, although other assessment practices such as the use of model answers and target-setting are also included.

Department priorities

The project also came out of concerns specific to our history department. My part 2 project was partly prompted by the 226 student responses to the 2013-14 department survey, which showed that our two lowest scoring areas in history were in response to the following statements:

- *Feedback on my work helped me clarify things I did not know and helped me improve the quality of my answer* (87% satisfaction rate)
- *Clear information about the assessment of this course and marking criteria was provided* (88% satisfaction rate).

Despite increases in both areas in the following 2014-15 survey, feedback remained the department's lowest scoring area, and therefore merited continued research. The challenge of applying new specification changes to the department this year also presented an opportunity to review our current assessment processes, allowing for the creation of new schemes of work that could be built around a greater focus on formative assessment.

College priorities

Key college priorities this year have consisted of 'studentship' and using Teacher Learning Communities (TLCs) to explore ways to develop student independence and resilience. Teacher Learning Communities are described by Skerrett as 'those that continuously inquire into their practice' in order to 'negotiate new meanings that improve their practice' (Skerrett, 2010: 648). In our college, Teacher Learning Communities were only established last year and are still therefore developing into an integrated part of teacher professional development. My research project contributed considerably to the work of our TLC by helping to provide a focus

and a research base for other colleagues to use in their own investigations and the project has helped to more effectively embed our TLC by providing this research backing. Thus, my research not only contributed to the wider college aim of establishing research-driven TLC's, but the focus on formative assessment also contributed to the college priorities of developing studentship.

National priorities

Concerns over current assessment practices, but specifically teacher marking, have been highlighted this year, with marking often being cited as a key reason for workplace stress. For example, a report by the Independent Teacher Workload Review Group from the Department for Education entitled 'Eliminating unnecessary workload around marking'[1] reported in March 2016 that the current marking burden on teachers 'must be addressed' and outlined an 'aim of shrinking the importance marking has gained over other forms of feedback'. This message was reinforced by a report from the NASUWT at their annual conference on the 26th March, which stated that 'Marking and assessment policies were cited by over three quarters (76%) of teachers in the NASUWT's annual Big Question survey as the biggest generator of excessive workload in their school'[2]. Developing student-centered formative assessment practices and mindsets could play a role in reducing the teacher's role in assessment and therefore potentially contributing to a reduction in teacher workload, as has been highlighted by researchers such as Price et al. (2010) and Boud and Molloy (2013).

[1] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/511256/Eliminating-unnecessary-workload-around-marking.pdf [Accessed 3 April 2016]
[2] http://www.nasuwt.org.uk/Whatsnew/NASUWTNews/PressReleases/NASUWT_015486 [Accessed 3 April 2016]

If teachers are to play a reduced role in the assessment process, this would not only be to the benefit of teachers. Empowering the student voice in the assessment process should benefit students, as a more inclusive and democratic process would encourage greater student participation, and this could be extended to student participation in wider public life. Torrance and Pryor see classroom assessment as 'a set of social practices' and 'a site of immense significance with respect to … how children learn about themselves' (1998: 20). It would be an overstatement to stretch Torrance and Pryor's analysis too far, and to say that formative assessment therefore plays a significant role in preparing students to engage and participate in public life outside of school, but if schools are a microcosm of wider society where students learn how to become citizens, it might be important to consider the messages that some of our current teacher-dominated assessment processes are giving to students about the importance of their voice in determining outcomes. Unease following Britain's EU referendum has pitted younger "Remain" voters against older "Leave" voters, but the greater issue worth exploring might be why it has been estimated that 64% of young voters (aged 18-24) didn't vote at all[3]. Arguably, if this issue is not being tackled in classrooms across the country -as the key forum for communicating with young people- it is unlikely to be seriously tackled anywhere. These concerns about encouraging student participation and democratic classrooms not only influenced my decision to continue studying formative assessment, but also influenced the methodological design.

In sum the rationale for this project is located within concentric spheres of influence: immediately, it responds to concerns regarding my own professional practice; more widely it is set within departmental and college priorities; it can then be placed within contemporary

[3] http://www.independent.co.uk/voices/eu-referendum-brexit-young-people-upset-by-the-outcome-of-the-eu-referendum-why-didnt-you-vote-a7105396.html [Accessed 3 July 2016]

research literature that focuses specifically on the implementation of formative assessment; and finally can be seen to be broadly located within wider political concerns about teacher workload and the apparent 'crisis of motivation amongst young people' (Ecclestone, 2002: 2).

Definitions

To clarify some key definitions used in this assignment:

- "Assessment policy", as it applies to our department, is not only a generic mission statement, but also consists of specifically ascribed assessment processes, including the regularity of teacher marking and the types of formative assessment used in the classroom to ensure a degree of parity across the department.
- "Formative assessment" is a broad term, and arguably covers any activity which helps students learn from assessment. My primary focus in this project has been on self rather than peer assessment of written work, but other assessment methods such as the use of model answers and target-setting have also been included.

Literature review

The starting point of my project was to complete a review of the relevant literature. This was an essential first step to inform my thinking, and then to share this with colleagues to develop the department's assessment policy for this year. It also allowed me to further explore some of the questions raised by my part 2 project, particularly some of the concerns highlighted by the students in focus groups, and I have taken quotes from these focus groups to provide the starting point to each section of this literature review.

To borrow the language used by some History educationalists of “big pictures” and “little pictures” to explore the different scales of history (e.g. by Nuttall, 2013), this literature review, and more widely this whole project, sits concurrently within two different perspectives. Firstly, there are the “big pictures”, or wider contexts of formative assessment, which provide the justification for pursuing it, and then there are the “smaller pictures”, which consist of the specific implementation and application of formative assessment in the classroom. These “pictures” are then subdivided into further categories:

- The “big pictures”: theories of learning and alignment; democracy and rights; the role of motivation
- The “smaller pictures”: measuring the effectiveness of formative assessment; the feedback loop; the relative roles of student and teacher in the feedback process; applying formative assessment in the classroom.

Both of these pictures were necessary to developing a meaningful formative assessment policy. I found that having a secure knowledge of the theory behind formative assessment and its wider context was hugely important in allowing me to develop meaningful formative assessment tasks in the classroom, and in presenting these tasks to the students in a more convincing and assured way. It seems my previously weak handle on the theory behind formative assessment is shared by others - as Dann highlights, ‘The area of self-assessment is one which is often promoted but little understood’ (2002: 73).

Equally, however, an understanding of the theory offers little to the practitioner without some thinking about how it can be implemented, and extricating the differences between the two is challenging and perhaps unrewarding. Indeed, formative assessment tasks themselves have been described as a ‘Trojan horse’ (Black et al. 2006: 127) and that what is really being

measured is the change in values and priorities that the activities encourage, rather than strictly a measure of the efficacy of the formative assessment tasks themselves. Therefore, the division of my literature review is for clarity, but it is perhaps necessary to use Claxton's 'split screen' thinking (2007: 122) to ensure both pictures are simultaneously kept in view for the practice of formative assessment to be meaningful.

The "big pictures"

Developing a formative assessment policy in our history department was never an end in itself. As Ecclestone states, 'a deeper commitment lies behind these goals' (2002: 12). It therefore seems necessary to start the literature review with these wider commitments, before focusing on the specifics of formative assessment.

Theories of learning: aligning the classroom context with formative assessment

'We are told from an early age that teachers are like...in charge'
- "Susanna", part 2 focus group.

Formative assessment does not take place in a vacuum – or at least it shouldn't. Black states that 'formative assessment ought to be intimately connected with the processes of teaching and learning' (1993: 51) and Black and Wiliam later reiterated that, 'underlying the various approaches are assumptions about the psychology of learning' (1998: 16). Many would agree that this 'psychology of learning' broadly falls within constructivist theories of learning; a link made explicit by Black and Wiliam, who highlight the importance of applying self-assessment 'in the context of a constructivist classroom' (1998: 10). The extent to which Black and

Wiliam's models of formative assessment are truly constructivist is debatable; nonetheless, Black and Wiliam are certainly clear that models of learning provide an essential context to formative assessment, stating that assessment approaches 'interact strongly with the pedagogy adopted' (1998: 39), and they raise the important point that these models of learning are held by both staff and students- 'students bring to their work models of learning which may be an obstacle to their own learning' (1998: 30). This is perhaps best put by Dann, who states that, 'If pupils are not encouraged to participate within their learning environment…then their involvement in self-assessment is unlikely to be of value' (2002: 127). Alignment of pedagogy and formative assessment therefore seems to be key to the success of formative assessment; a point further reiterated by Stobart, who argues that 'for formative assessment to lead to learning, the classroom context has to be supportive' (2012: 236). The case seems to have been made then for the importance of placing formative assessment within a constructivist context, but it is worth noting that even the assumed benefits of the constructivist classroom could be challenged. Torrance and Pryor highlight the 'social and cultural context of the classroom, where individual pupils have different access to power, means that the more collaborative approach needed for divergent assessment is not always equally (and equitably) available to all pupils' (1998: 155); although not explicitly stated here, the constructivist classroom arguably suits the extroverted learner, and rewards those who can more confidently navigate the social space of the classroom. Nonetheless, it certainly seems sensible that assessment processes that emphasize the centrality of the student will be most effectively embedded in theories of learning that are also student-centered, despite legitimate concerns about the extrovert bias in constructivist classrooms.

The context of the constructivist classroom is perhaps even more important when it is considered that formative assessment should not be seen as an end in itself, nor necessarily a

guarantee of student independence. Ecclestone argues that in the wrong contexts, formative assessment can become 'a mere "technology" where practices are divorced from social and educational aims and become, instead, impoverished ends in a compliant, ultimately meaningless, pursuit of performance targets' (2002: 13). As Black and Wiliam state, there is a 'close link of formative assessment practice…with other components of a teacher's own pedagogy' (1998: 20). However, an argument has been made that this interconnectedness of assessment and pedagogy has been taken too far. Torrance and Pryor have argued that 'teachers may be better advised to think of formative assessment as part of their pedagogy' (1998: 152), a statement which seems to shift the responsibility for formative assessment from the student to the teacher. Wiliam's theory of assessment expressed in his recent book on embedded formative assessment continues this theme (2011). His statement that 'assessment is *the* central process in instruction' (2011: 47) gives credence to the oft-made criticism that assessment for learning is giving way to assessment *as* learning (as made by Torrance, 2007: 281). Wiliam's language of 'instruction', almost suggests a return to behaviourist models of learning where the teacher 'instructs' the students in assessment criteria, and seems to contradict their earlier claim that 'a commitment to the use of formative assessment necessarily entails a move away from unitary notions of intelligence' (1998: 56). This tension has arguably always been implicit in Black and Wiliam's work, and particularly their feedback loop model which will be discussed later in this review. It would therefore seem that formative assessment cannot automatically be seen as constructivist where it's application consists of teacher-transmitted assessment criteria, as there is little room here for the student construction of knowledge.

The alignment of formative assessment to the system of national tests

> *'…but on results day, I won't get a piece of paper with feedback comments on it, I will get a piece of paper with three letters on it…At the end of fourteen years of education you don't get "you did these three things well but this could have been better, try again next time". That's it'*
> - "Susanna", part 2 focus group.

Torrance and Pryor identified in teachers' self-reported accounts 'two discourses running simultaneously – one about child-centered schooling and the teachers' motivation to listen to, get to know, and teach young children, and a second concerned with measurement, categorization and accountability'; as they then state, 'Reconciling these two accounts of the purpose and practice of "teacher assessment" was proving extremely difficult' (1998: 42). For sixth-form teachers (although not exclusively), this second discourse can dominate teaching because of the system of national tests. As Susanna had identified (in the above quote), a key problem with formative assessment was that it did not seem to her to reflect how national assessments were conducted, and therefore it felt less relevant- the two did not align; she too had identified Torrance and Pryor's clashing discourses of the constructivist classroom with the system of national assessment.

The case could however be made that the shift away from modular assessment has made the process of reconciling these two discourses easier. Sadler asserts that modular assessment is potentially damaging to meaningful formative assessment because it replaces the 'longer-term goal of excellence' with the 'drive to accumulate credit' - a position Sadler justifies by stating that modular courses 'produce in students the mindset that if a piece of work does not contribute towards the total, it is not worth doing', and further, 'the length of each unit is often not long enough…there is simply not the time to do it' (1989: 141-2). Therefore, with

the changed A level specification to linear assessment, arguably there is now greater scope to apply meaningful formative assessment as Sadler understands it, rather than more rigidly "task compliance", potentially providing greater alignment between the nationwide assessment system and formative assessment. The point is furthered by Torrance and Pryor, who claim that it is possible that there is so much summative assessment at a national level that teachers are driven to do more of it in the classroom (1998: 11), and so by extension less summative assessment provided at a national level by linear courses might allow for more "breathing space" for our department to better embed formative assessment practices at the planning stage of a new specification.

Democracy and rights

'Yeah, they're judging you'
– "Alan", part 2 focus group.

Wider still than the educational and assessment settings, are concerns regarding democratic schools and the rights of students. Again, arguably this is an issue of alignment; formative assessment that is aligned with a democratic classroom is likely to be more effective because they are mutually reinforcing. More importantly however, democracy and student rights are one of Ecclestone's 'deeper' commitments (2002: 12) that drive the need for formative assessment in the first place.

Firstly, for the alignment concern: Hattie argues that for formative assessment to be effective, 'students need to feel that they "belong" in learning' and that 'there is a high level of trust' within the classroom (2012: 121). "Belonging" can still be hierarchical – patriarchal even- and

Hayward takes this further to argue for the 'right of learners to have their voices heard' (2012: 137). Formative assessment is perhaps better based in the context of this latter position. Not only is this for practical reasons- as Dann states, 'Trying to ensure that pupils are encouraged to articulate the basis for their judgements… as part of the self-assessment processes' (2002: 135) clearly helps the teacher understand the current level of student understanding, but furthermore, an important part of making formative assessment meaningful for students will be ensuring it takes place in a context where pupils' ideas are valued (Dann, 2002: 127).

Secondly, the interest in student rights as a 'deeper' commitment: Hargreaves argues that assessment for learning involves students 'taking some control' (2005: 213). Therefore, it could be taken at face value that formative assessment empowers the student voice in the assessment process and therefore gives the students greater rights in the classroom and helps develop a more democratic environment. However, this is further complicated by two key issues. Firstly, although formative assessment sets up assessment as dialogue rather than transmission, as Black highlights, in that dialogue, the 'teachers' power easily overwhelms' (1993: 81), so the process itself is no guarantee that "control" has been transferred. Secondly, Torrance highlights that in practice, much classroom formative assessment really consists of teaching 'criteria compliance' and this risks making students even more reliant and dependent on their teachers, as it is they who hold the 'key' to understanding how to write an essay (2007: 282).

Even without these concerns, the very premise of giving voice to students through formative assessment can be challenged. In the classroom environment, students are not seriously offered a choice as to whether or not they participate – the formative assessment activity is

set by the teacher and it is therefore a requirement that the student participates. Dann questions 'the rights of teachers…to demand or expect the pupil to reveal his/her feelings and reflections on learning' and highlights '...the possibility that more direct involvement of pupils in processes of self-assessment may only serve to increase and deepen forms of surveillance' (2002: 75). However, this can to an extent be reconciled by restoring Sadler's focus on making the student central to the assessment process, so that they 'eventually become independent and fully self-monitoring' (1989: 120). By encouraging students to measure their own progress, formative assessment becomes less a process of teacher intrusion and more one of developing student skills in self-assessment.

In sum as Hayward argues, 'assessment for learning is radical' (2012: 126), and for it to be implemented effectively it requires a radical shift in the balance of power within the classroom. Clarke reinforces the point; 'the ability of teachers to shift the locus of control in the classroom from teacher to pupil – to "let go" – is often the attitude change needed' (2008: 2). It is not however an uncomplicated process, and reviewing this literature has not allowed me to draw any concrete "solutions" to this problem. It has however encouraged me to constantly reflect within my practice on how I am asserting power within the classroom and how I am giving space to student voice. I hope that at least this engagement with the issue will to some extent contribute to creating Clarke's 'attitude change' in the classroom (2008: 2).

Motivation

'I don't think it's effort, I think it's a confidence issue'
- "Ruby", part 2 focus group.

Clarke argues that formative assessment 'promotes confidence that every pupil can improve' (2008: 10). This can be justified in so much as it empowers students to take control of their own learning, but perhaps in itself is a little simplistic. Formative assessment can also be stressful, and Torrance and Pryor perhaps provide the more nuanced analysis in stating that the impact on 'learning and motivation…might be negative as well as positive' (1998: 131). It is therefore perhaps not the case that formative assessment is inherently motivating simply because it places the student at the center of the process. For example, Stobart has argued that, 'the emotional and effort costs of acting upon it [feedback] may be too much, particularly if there is low commitment to it' (2012: 240-1). Although Stobart was writing about responses to teacher feedback here, by extension this analysis can also apply to self-assessment – in order for students to really gain from it they need to be prepared to negotiate and accept it, and this requires student motivation. An effective formative assessment policy therefore ought to not only consist of classroom practices, but also consider how to motivate students to engage with them.

Black and Wiliam's literature review highlights the finding that 'task-involving evaluation is more effective than ego-involving evaluation' (1998: 13), a point reiterated by Wiliam in 2011 (2011: 119), thus it seems the case that feedback that focuses on targets and the work itself rather than the individual student is the more motivating. However, feedback which removes the student from the process must necessarily focus on the product – and yet this risks only reinforcing a student's view of assessment as performance – an attribute highlighted by Harlen as being characteristic of extrinsically motivated students (2012: 176). Whilst teacher

feedback can focus on setting specific *methods* to improve, shifting the focus from product to process, much formative assessment consists of students making *judgements* about student work (leading Taras to conclude that all self and peer assessment is in fact summative, 2009: 64). Self and peer feedback therefore might reinforce views of assessment as being about the criticism of the product (and by extension the producer) rather than evaluation to learn from, leading us to the conclusion that some formative assessment practices risk reinforcing the view that assessment is about a product rather than a process. Nonetheless, an essential part of undermining ego-orientation consists of breaking down student beliefs in fixed ability (as claimed by Galloway et al., 2004: 99; and Claxton, 2009: 179), which is also an important method of motivating students to take part in formative assessment (which is arguably built on an assumption that all students can improve and ability is therefore not fixed). Therefore, if formative assessment risks a greater focus on the product, undermining beliefs in fixed ability might mitigate the extent to which this is taken as a reflection on the producer.

Sharing the aims of formative assessment with students might also be an important way of motivating them. Simply sharing the theory behind formative assessment with students might increase trust in the process. I have used an analogy employed by Clarke this year to challenge the way my students think about assessment. Clarke compares learning to a driving test, pointing out that very few fail their first test and give up, they just get more lessons and try again until they pass because they are motivated to drive (2005: 20). It would be difficult to assess the extent to which the sharing of aims and theory does increase motivation, but educationalists such as Claxton have certainly claimed its significance (2007: 128).

The "smaller pictures"

Measuring the effectiveness of formative assessment

> *'I know the aim is to make it so that our writing does get better but when you're not confident with it, it just doesn't really have that effect'*
> - "Ellen", part 2 focus group.

An important part of developing our department's formative assessment policy would consist of establishing some success criteria to establish what was and was not working within the department's current formative assessment practices in order to provide some direction for development. The literature on measuring the effectiveness of formative assessment is fairly unanimous in its conclusion that this is a tricky task. As Wiliam states, 'It cannot be proved that formative assessment did improve learning as that would require the counterclaim that what happened would have been different to what would have otherwise happened' (2011: 43). That said, the educationalists perhaps most associated with attempting to "prove" the effectiveness of formative assessment are Black and Wiliam. In 'Assessment and Classroom Learning' (1998), their review of experiments conducted with control groups suggest the value of self-assessment methods in improving student performance as measured through pre and post test data, and they cite studies such as Schunk & Rice, (1991) and Schunk and Swartz (1993a) who used not only learning performance but also student beliefs about their own performance as measures of effectiveness (1998: 23). However, Bennett has argued that the 'pooling of results' from such 'disparate' studies does not allow for a meaningful meta-analysis (2011: 11). Concerns with the approaches taken within specific studies are also highlighted by Black and Wiliam themselves, and they cite Bangert-Drowns et al (1991a) in considering the difficulty in using pre-testing to measure effectiveness as it can give learners practice in, or act as advance organizers of, the material to be covered (1998: 51), and their

point that assessment results may be better indicators of motivation and task completion than of student understanding further undermines the role of pre and post testing data if instead what has really been measured is task preparation (1998: 57). Indeed, Black and Wiliam raise a number of caveats to their conclusions, including their claims that, 'it might be difficult to separate out the particular contribution of the formative feedback to any learning gains' and 'the demand for unambiguous quantitative comparisons of effectiveness can never be fully satisfied' (1998: 16). It seems surprising therefore that despite this relatively cautious treatment of proving efficacy, Black and Wiliam ultimately conclude that, 'The research reported here shows conclusively that formative assessment does improve learning.' (1998: 61). Indeed, despite the caution issued about 'unambiguous quantitative comparisons of effectiveness', in a later collaboration Wiliam does place a figure on this, and claims that the 'use of formative assessment can increase the rate of student learning by somewhere between 50 and 100%...' (Leahy and Wiliam, 2012: 52).

Ultimately, many studies that have attempted to quantify the improvement made by students using formative assessment contain some methodological flaws because of the impossibility of controlling for all variables in real school settings. These studies still have a use in exploring the value of formative assessment, but where universal conclusions are necessarily so difficult to arrive at, it is overdrawn for educationalists to claim 'proof' of formative assessment's efficacy, which arguably undermines the purpose of pursuing such research in the first place. It is perhaps more powerful to cite examples of where it has worked in local contexts and to then encourage practitioners to explore this for themselves.

Dann, for example, adopts a methodological approach that is more removed from attempting to “prove” the effectiveness of formative assessment by instead asking pupils whether they thought that self-assessment had helped them to improve and then exploring meanings behind their responses (2002: 91). This was an approach also adopted by Hanrahan and Isaacs (2001) who sought to understand how students felt about self and peer marking and then use these views to further ‘develop assessment techniques that maintain the perceived positives and limit the perceived negatives’ (2001: 54). Dann found that ‘although pupils were positive about the potential to help improvement, most were not able to give specific details’ (2002: 91). Certainly, asking pupils to identify what they thought was effective is not a measure of effectiveness in itself, and results might lack reliability if students think that to agree is the “right” answer, which might in this case be particularly indicated by the fact that many of Dann’s pupils could not justify why they felt it helped them improve (2002: 91). Nonetheless, student perceptions of efficacy are one measure of effectiveness, as presumably if a student considers something to be effective they are more likely to engage with it and thus its impact would be enhanced. The process of exploring that question also allows the researcher to better understand how students think about their role in the assessment process which would also provide the researcher with an insight into its effectiveness in their own local context. It is this ‘discovery’ approach (Cohen et al., 2000: 53) towards research that has had the greater influence on this project compared to a positivist scientific approach in measuring effectiveness.

The feedback loop

'I need to jump through hoops'
- "Susanna", part 2 focus group

Perhaps the most dominant model of improvement used in formative assessment is that of the feedback loop. The language of feedback loops was used by Sadler (1989: 120) and then further developed by Black and Wiliam (1998). The model provides direction for student improvement whereby a "gap" is identified between a student's current performance and the desired level. Specific "inputs" are then provided in order to close that gap. Although the model has been criticized for being too reductionist, both Sadler and Black and Wiliam accept these criticisms. Sadler describes the feedback loop as making 'programmatic decisions' (1989: 120) but then also highlights the tension between using such a mechanical model in a process dependent on qualitative judgments (1989: 139) and points out that 'student development is multidimensional rather than sequential' (1989: 123). Black and Wiliam state that it 'would be a mistake to regard the student as the passive recipient of a call to action' (1998: 21) and they highlight the 'complex links between the way in which the message is received…and the learning activities which may or may not follow' (1998: 21). Indeed, as Dann highlights, it is difficult to know when students have mastered a concept (and thus closed the loop), and just because they can demonstrate it in one scenario does not mean they can automatically do it again (2002: 56). Hargreaves is particularly critical of Black and Wiliam's model of the feedback loop, arguing that the model of closing the gap and moving forward lends itself to a theory of 'assessment as measurement' (2005: 220), and in a critique more specifically targeted at Black and Wiliam, she argues that the 'black box represents the version characterized primarily by a measurement/objectives conception of assessment and learning' (2005: 223). Hargreaves' critique is perhaps necessary in order to challenge an overly

mechanistic interpretation of the feedback loop model, but is perhaps overall a little unfair. It is important to point out that the feedback loop is only a model, and should not be mistaken for an exact replica of the feedback process. It does however highlight an important point about who is responsible for closing the feedback loop.

The relative roles of student and teacher in the feedback process

'…teachers are like the word of god to me'
– "Anita", part 2 focus group

Sadler's point that, 'Teachers use feedback to make programmatic decisions with respect to readiness, diagnosis and remediation. Students use it to monitor the strengths and weaknesses of their performances' (1989: 120-1) highlights the issue of whether or not "closing the feedback loop" is the teacher's or the student's responsibility, and raises a wider question about the relative roles of student and teacher in the feedback process. For example, Boud and Molloy (2013) discuss two models of feedback; one where the teacher drives the feedback -a model which they situate within the disciplines of engineering and biology- and an alternative model where the student has a 'key role in driving learning', which they describe as 'sustainable assessment' (2013: 698). They perhaps unfairly place Sadler within the first model (2013: 702), which although can be seen in Sadler's view that the teacher should provide structure and guidance to the feedback process, it perhaps overlooks Sadler's ultimate aim – to help students to become self-monitoring. This interpretation of Sadler's work is also challenged by Taras, who argues that Sadler's feedback model is more focused on student responsibility, whilst the work of Black et al. (2003) 'weight[s] the responsibility firmly onto the teacher' (Taras, 2009: 61). Certainly Wiliam's more recent work can perhaps be more comfortably located within this model. His emphasis on the teacher's

responsibility for the feedback loop is evident in his book *Embedded Formative Assessment* where he argues that during the feedback process, 'in many cases, the decisions will be made by the teacher' (2011: 43) and that the purpose of assessment is to 'inform the teacher what needs to be done next' (2011: 77). Indeed, the overall focus of his work is on teacher action and teacher responsibility, and although he concludes that 'feedback should be more work for the recipient than the donor' (2011: 129), in the context of this book this statement feels a little contradictory. As Nicol states, 'Producing feedback is more cognitively demanding than just receiving it' (2010: 514), thus suggesting that in Wiliam's model the teacher is engaged in the harder cognitive task than the student. Therefore, taking Nicol's point into account, the model of self-regulation that Sadler proposes is perhaps preferable in ensuring it is the student who is engaged in the 'cognitively demanding' part of the feedback process (2010: 514).

However, despite the weight given by many researchers since Sadler (e.g. by Nicol and Macfarlane-Dick, 2006; Blair and McGinty, 2013) to the role of the student in the feedback process, the role of the teacher cannot and should not be erased entirely. Dann argues that in assessment conversations, the teacher 'must remain as the senior partner' (2002: 102) and this reflects Sadler's view of recognizing the teacher as expert in the feedback process (1989: 127), which a purely constructivist model might underplay. Nonetheless, this can again be complicated by considering the twin aims of feedback – reliability and utility. Certainly, teachers have a clearer understanding of 'quality', and thus can provide the more reliable feedback. However, just because it is reliable this does not guarantee that it is useful to student learning. Student self-assessment therefore arguably provides a more useful method of assessment that encourages students to learn from the process of assessment, although at the expense of a judgement that is as accurate as the teacher's. If formative assessment by

definition prioritizes utility while summative assessment prioritizes reliability (Harlen and James, 1997: 371), then teacher expertise is to a small extent less significant to formative assessment because arguably their expertise is more relevant to summative judgements. Thus, to an extent, to see the teacher as a senior partner in the formative feedback process arguably misconstrues formative assessment as being a process that needs to be reliable rather than useful. Teachers nonetheless clearly have an important role in sharing their understanding of "quality" with the student, in order for the student to then take responsibility for their own progress, but much like the concept of "scaffolding", this should perhaps be gradually reduced over time.

Applying formative assessment in the classroom*.*

'It's because the grade is the only thing that matters'
- "Alan", part 2 focus group.

Black and Wiliam argue that 'there is clearly no single royal road' in applying formative assessment (1998: 61), a point they followed up in 2012 by reiterating that developing formative assessment practices is not about developing 'a recipe for teachers to follow' (2012: 15). As Leahy and Wiliam argue, teachers need to be able to modify formative assessment to fit their specific classroom needs rather than attempt to import strategies wholesale (2012: 56). That said, there are overarching principles and ideals that are necessary to effectively embed formative assessment in the classroom. Torrance and Pryor perhaps offer the best characterization of an approach to developing formative assessment in their claim that, 'formative classroom assessment can never be reduced to a set of procedures or practices that will "work", but rather should be seen as an open, interactive process that might "get somewhere"' (1998: 159).

If there is a reticence in the literature to provide a "recipe" for formative assessment policies, there is greater clarity regarding what formative assessment tasks should not look like and the overarching principles that should be aspired to. Torrance argues that much formative assessment in the post-secondary sector has 'an overwhelming focus on criteria compliance' (2007: 282). Black reinforces this by pointing out that self-assessment can risk becoming mechanistic, and so efforts should be made to ensure self-assessment activities are genuinely thoughtful (1993: 81). In order to achieve this genuine thoughtfulness, Dann suggests that, 'A conscious effort needs to be paid to ensuring that pupils have appropriate opportunities to express the factors which have influenced their work' (2002: 138). This is furthered by Torrance and Pryor (1998) who highlight the importance of asking genuine questions in order to empower the student -which has significance for formative assessment because the teacher does not know which factors 'have influenced their work' (Dann, 2002: 138)- and therefore when the teacher engages in a discussion with the student about their work it takes place on a more even footing compared to the more 'ritualistic' 'classroom performance' of much teacher questioning where the teacher already knows the answer (1998: 151-2). The process of 'asking a pupil to clarify' encourages a 'process of reflection' which can then 'put the pupil in a position of relative power because the pupil is likely to know the answer in a way the teacher doesn't' (Torrance and Pryor, 1998: 163). The importance of encouraging this 'dialogue' is specifically cited by many researchers, including Nicol (2010), Carless et al. (2011), and Orsmond et al. (2013), therefore it seems clear that however specific formative assessment activities might be developed in the classroom, the notion of dialogue ought to be central to their practice.

Linking the literature to classroom practice

This literature review provided a direction for the development of our department's formative assessment policy, and some key aims were established at the start of the academic year. Increasingly challenging formative assessment tasks were built into schemes of work, so that students started with simpler tasks of highlighting key criteria and moved towards writing fuller evaluations of their own work, to attempt to gradually increase student independence and to move towards genuinely 'thoughtful' self-assessment (Black, 1993: 81). In order to establish a dialogue in the feedback process, it was decided that students would add their own self-marked comments and questions to their work before teacher submission, in order for teacher feedback to specifically respond to those concerns. The value and purposes of formative assessment was also shared with students to help them 'become insiders rather than consumers' (Sadler, 1989: 135).

Reviewing the literature also allowed me to consider different measures of effectiveness in order to develop a methodology that could explore then effectiveness of these changes in order to make further improvements.

Methodology

Research questions

My aim to use student views to develop our department's formative assessment policy required some consideration of what was currently working, alongside building up a picture of how the students experienced the policy in order to make further adaptations and changes. The focus here is on the *development* of an existing policy rather than the creation of something entirely new, thus there is a focus on both establishing what was "working" from existing practice, and using these insights to then determine how to improve the policy for future use. Each research question explores a different definition of effectiveness because as Cohen et al. emphasize, the importance of cross-referencing different types of data and asking a range of questions is particularly important when exploring more complex conceptual thinking – something they describe as using more 'slices of data' in order to build a fuller understanding (2000: 76).

The research questions that developed out of my rationale and literature review were the following:

1. Do students report more engagement with self-assessment compared to the students that I taught last year? (Questionnaire)
2. To what extent do students consider current self-assessment strategies to be effective? (Questionnaire and group interviews)
3. To what extent do students report feelings of responsibility within the assessment process? (Questionnaire and group interviews)

Thus, these three research questions each provide a different measure of effectiveness:

- The first research question provides some indication of the success of the department's formative assessment policy this year by comparing it with the "baseline data" from last year's students.
- The second question explores student views on the actual practices of the assessment policy itself – Black and Wiliam's 'kit of parts' (1998: 38)- to explore their effectiveness, and to some extent reflects the methodological approach taken by Dann (2002: 91) (as previously discussed).
- The third question locates student views on assessment practices within their wider understanding of the role of the student in the assessment process. This addresses a key research aim, as formative assessment is not in this project being pursued for its own sake, but in order to develop student independence and responsibility; therefore, for a valid measure of success, it is necessary to also explore these attitudes.

In isolation, none of these questions definitely "measure" the success of the department's formative assessment policy this year, and this was not the primary aim of the project. Instead, this project can be located within the 'discovery perspective' as defined by Cohen et al., where a researcher aims to 'gain understanding of how individuals make sense of their worlds' (2000: 43). For example, the second research question explores student views of specific formative assessment practices, but this should be seen less in the context of Bennett's call to isolate formative assessment practices and then apply a 'scientific' 'standard of rigour' to measure their individual effectiveness (2011: 15), but rather fits within Swaffield's call for research to 'check constantly for the actual (as opposed to the intended) effects of practices' (2011: 438). Nonetheless, the use of baseline data in response to the first

research question does provide some attempt at "measuring" success to place the more interpretivist approach of the second and third research questions into perspective – an approach described by Cohen et al. as offering 'complementary insights' to strengthen conclusions (2000: 53).

These research questions were explored through the use of questionnaires and group interviews. Questionnaires were chosen because they provided every student with an opportunity to take part in the research project, which focus groups could not do. Group interviews were also used to explore the second and third research questions because of their more discursive nature, allowing students more space to develop and express their views than the necessarily more restrictive format of a questionnaire.

Planned data collection

My sample consisted of three AS History classes. These students were selected instead of A2 History students because the AS students were starting a new specification, so a new formative assessment policy could be more easily embedded into a new scheme of work. Additionally, working with AS students meant that they would be able to see the changes made to the policy as a result of their views next year, with the intention that this would make the project more meaningful and useful to them. The timing of the data collection was as follows:

Questionnaire 1	March 2016
Questionnaire 2, Questionnaire 3 pilot	June 2016
Questionnaire 3	July 2016
Group interviews	July 2016

Questionnaire 3 was piloted with an AS Politics class. The class raised no concerns with the questions themselves, but did suggest changes to formatting to make the questionnaire clearer, which I adopted. Questionnaires 1 and 2 were not piloted. Questionnaire 1 had been piloted last year, and for the purposes of validity I did not make any changes to the questioning. Questionnaire 2 proved too difficult to pilot as it asked questions specific to practices being used in the History department.

The group interviews were conducted by a colleague who also has an interest in formative assessment, but does not work within the department. This was an attempt to reduce researcher bias, and also to further share the research process with other colleagues. Group interviews and questionnaires were all conducted during class time.

Process of data collection

The data collection timetable was kept to, although an issue arose with the timing of questionnaire 2, which had a relatively low return rate of 73% due to a number of students being absent from lessons that week due to a college trip. Furthermore, the absence of a pilot for questionnaire 2 due to the specific nature of the questioning meant that I had not anticipated that some students might struggle with the rank ordering questions, and where they were unsure of an order, would put multiple options in the same rank. This made my chosen method of data analysis – to ascribe a numerical value to each rank in order to provide an overview of preference – very difficult for those students, and in order to ensure clarity in my data analysis I chose to leave those questionnaires out of my calculations.

Ethics

Student participation: Because one of the key aims of this project was to develop student ownership over the assessment process, it was important that student voice was an integral part of the research process. A methodological model where I "mined" data from passive students would not have been in-keeping with my aim to develop a project that was student-centred and democratic. I therefore aimed to design a process of data collection that would be beneficial to the students, and to treat them as participants in research rather than as subjects of research. Not only this, as my literature review indicated, a wider justification for developing formative assessment was to challenge the top-down and authoritarian view of the classroom in favour of a more democratic one. Therefore, to be true to this aim, my research needed to be accessible and democratic. For these reasons, I chose to complete all of my questionnaires on paper and in the classroom, providing equality of access. All questionnaires were anonymous to allow students to express themselves freely. Whereas in my part 2 project I had used focus groups, and found that these had provided my richest data set, I was aware that the more "school friendly" students would be those most likely to attend. Therefore, for this project I held group interviews with one of my classes within lesson time, and these were conducted by a colleague. All students in the class were invited to attend one of three groups of 3-5 students, but were also given the offer to abstain, of which four students did. Whilst the interviews took place, the students remaining in the class were engaged in a discussion about the use of formative assessment in history this year, and were engaged in a group task to suggest improvements, and so would not feel disadvantaged by missing out on subject time.

Avoidance of control groups: Black and Wiliam's (1998) literature review cites the importance of studies that use control groups to judge the extent of student improvement as a result of

using self-assessment methods. However, this would not be a valid method for me as the use of control groups in my context as a classroom teacher would be unethical as it would provide an inconsistent experience to my different classes. Furthermore, as the research literature is so persuasive of the positive effects of formative assessment, I would likely risk disadvantaging some of my classes in the pursuit of attempted scientific accuracy – although the extent to which a control group genuinely isolates all other factors and thus can provide fair conclusions is certainly in doubt. There could have been an argument for different staff within the department trialing different methods, but again there was an issue of control, and the ethical consideration of students having different assessment experiences, which might only be further exacerbated in student eyes if a particular method was only associated with a particular member of staff.

Informed consent: Students in my classes were aware of and involved in my research project from the start of the year and thus were able to give more informed consent to their participation in questionnaires and group interviews. Questionnaires also each contained a 'tick box' at the end for students to show they actively gave consent to their data being used. At the start of the lesson where students were invited out to take part in group interviews, they were shown a list of indicative questions and the purposes of the interview were explained so that they could attend based on informed consent of what would be asked.

Data storage: All of the data collected was stored securely. Questionnaires were all anonymous, and stored in a locked filing cabinet. Group interviews were audio-recorded, and the recordings stored on a password-protected device. These recordings will be deleted after submission of my dissertation. Transcriptions have all used pseudonyms and were word

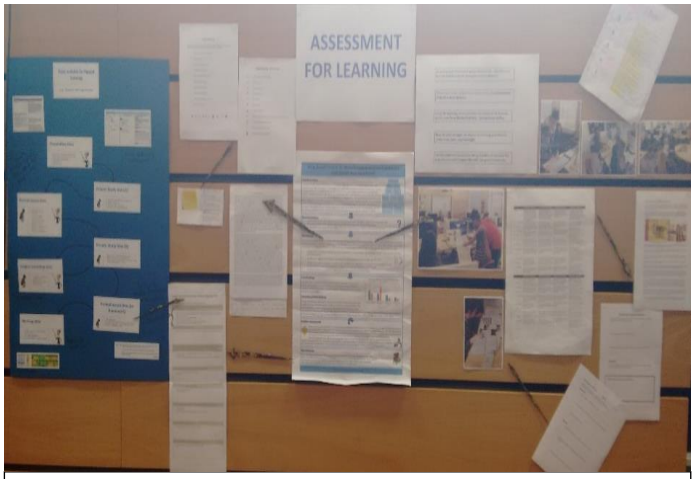
processed on a password-protected device. All names cited in this project are pseudonyms. All paperwork will be safely destroyed after completion of my dissertation.

BERA: My research was conducted within BERA research guidelines and I received ethical approval to conduct this study.

Collaboration

Within the history department: In the summer term, after my part 2 project was completed, we met as a department to discuss the implications of my findings, and to develop our new formative assessment policy for the next academic year. This was then further modified during an INSET on our first week back in September in light of my reading over summer. The impact of different features of the policy were regularly discussed and this meant that small details of the policy were adapted over the year in light of staff views.

Within the TLC: Collaboration outside of the department consisted of my work within the college's 'Teacher Learning Communities' (TLCs). After sharing my initial ideas with my TLC, the group decided to use my literature review and part 2 research to inform their own research projects. This ended up shaping our research direction for the whole academic year.



Poster produced for our TLC poster conference. My colleagues in the TLC chose to put the poster that I had produced for the MLT poster conference in the center to represent how their projects had been inspired by my research into self-assessment.

Many of my colleagues chose to focus on peer assessment as they felt this would be a more relevant application of my research in their department areas, and discussing their experiences gave me a useful contrast to my work on self-assessment. We shared our findings at regular intervals, and at the end of the academic year produced a "poster" to share our experiences and findings with the rest of the college.

My research could have been extended by including a formal collection and analysis of my colleagues' experiences of using formative assessment this year. However, early in the project I abandoned plans to include formal staff interviews as part of my data collection. This was initially due to pressures on staff time due to a colleague's long term illness in the first term, but also became a necessary decision in order to narrow the scope of my research. However, because policy-making in our department has always prioritized staff voices, I consider my research as important in formally prioritizing the student voice in this process. Nonetheless, a formal exploration of staff views concerning formative assessment would provide a valuable extension to this research project.

Process of data analysis

Questionnaire responses were all word processed to facilitate data analysis. Quantitative data was collected and transformed into percentages for the first questionnaire and charts for the second and third questionnaires to facilitate analysis. The use of percentages in analysing the data from the first questionnaire allowed for easier comparison to this baseline data. The sample sizes were also similar, thus making the percentages more comparable: my part 2 findings were based on 59 student responses compared to 58 student responses to the first questionnaire this year, and both were carried out at approximately the same point in the

academic year. Return rates of 98% for questionnaire 1, 73% for questionnaire 2 and 88% for questionnaire 3, providing samples of 58, 43 and 52 students respectively add validity to the findings of these questionnaires in terms of application to this local context. Rank ordering questions were processed by giving each rank a numerical value to present student preferences on a chart, again to facilitate analysis.

Qualitative data came in the form of open questionnaire responses and three group interviews. Questionnaire responses were first transcribed and then coded. The coding process was inductive and examples have been provided in the appendices. Although coding was inductive, many of the themes identified were also those identified by Hanrahan and Isaacs' (2001) study into the views of students who had recently undergone self or peer feedback, where student responses were coded within 'key dimensions' which included: 'difficult', 'gained better understanding', 'discomfort', 'productive', 'problems with implementation', 'reading other's work', 'empathy', and 'motivation' (2001: 53), thus providing a degree of validity to my coding in so much as these codes were also arrived at by other researchers in the field.

Group interviews were transcribed, and then presented alongside an analytical commentary (examples provided in appendices) in order to facilitate coding. These were then coded for key themes to support the cross-referencing of data with the questionnaire responses.

Findings

The findings presented here have been organized around the questionnaire data rather than delineated by research question. This is because an inductive approach was used to analyze the data whereby insights gained from any of the three questionnaires could be used to explore the three research questions to ensure insights would not be lost. This approach has been chosen over the decision to take a more positivist approach of organizing the findings to 'answer' each research question separately. As indicated in the methodology, because there is a particular focus on using the data to consider how students experienced the formative assessment policy, this inductive approach seems the most suited to this aim. Data from the group interviews are embedded into the second and third sections. The final summary of findings then returns to the research questions to provide some overall conclusions.

Questionnaire 1

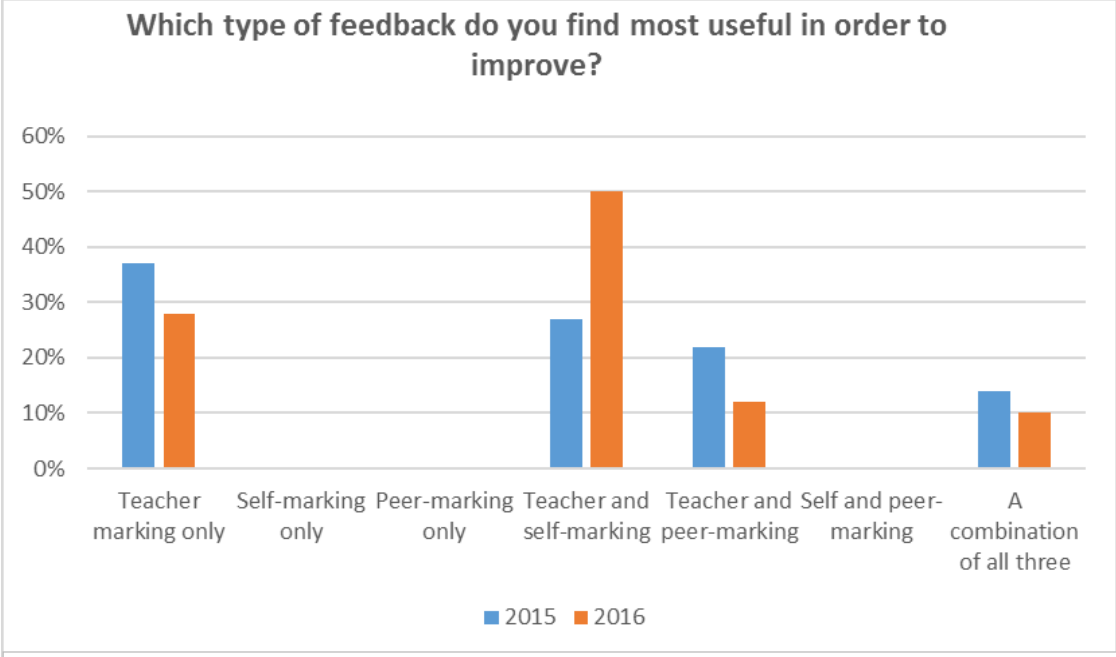


Figure 1: Student responses from questionnaire data gathered in March 2015 and March 2016 in response to the question 'Which type of feedback do you find most useful in order to improve?'

Last year, self-marking was applied in an ad hoc manner to some written assignments. This year, some form of self-marking was applied to every piece of written work. The data presented in Figure 1 does suggest increased engagement with self-marking this year, as the number of students claiming to find a combination of teacher and self-marking to be the most useful type of feedback to help them improve increased from 27% last year to 50% this year. It would not be possible to prove a causal link, as although the main change in policy was the regular application of self-assessment tasks, clearly the two years were not controlled for other variables. However, embedding self-marking into the assessment process may have raised its status as a "real" task – something emphasized by James (2008) as being important in motivating students. However, it might be difficult for students to really unpick the extent to which a type of assessment has helped them to improve, and thus this question may more accurately be measuring student motivation to engage with self-assessment rather than an accurate judgment of the extent to which it contributed to improved student outcomes.

Furthermore, the decline in preference for peer marking raises important questions about the extent to which student preference for self-marking has become a "learned response" due to repetition rather than evidence of fundamentally changed student values about the importance of students in the assessment process in so much as it could be considered that if students now prioritized the role of the student over the teacher in assessment, it could be expected this would also translate into increased preference for peer assessment. Alternatively, this could be a false measure of the extent to which the student role in assessment has become "embedded", and indeed the decline in preference for peer-marking is relatively small.

The extent to which this data does suggest greater engagement with self-assessment is complicated by the fact that no student selected 'self-marking only'. Indeed, although 12 out of the 28 responses to the 'teacher and self-marking' option focused on the advantages gained from the combination of both teacher and student feedback, many still gave slightly greater weight to the teacher voice, something also found in a study by Blair and McGinty, which they described as an 'expert-novice discourse' (2013: 471). A further four responses were separately coded for considerably prioritizing the teacher's voice over that of the student; for example, the comment 'Teachers are the experts, so know how to improve an essay or piece of work. Self-marking allows the student to see how they could have picked up extra marks' shows a recognition of some of the benefits of self-assessment, but clearly prioritizes the teacher's view. However, other views expressed a more shared approach towards assessment, where the teacher is instead placed into a supportive or "checking" role, such as the comment, 'Self-marking helps to show how to improve and teacher marking is good because the students feel more confident with the teacher's approval'. Comments such as, 'I can recognize mistakes I've made or what I haven't included, and I can also get a mark from my teacher to know if I'm on track' suggest a thoughtful understanding of the different roles teacher and self-assessment can serve (a level of reflection which in itself could be an indicator of genuine engagement), but this view was an exceptional one. Thus the data from this first question does overall indicate increased engagement with self-marking, but generally only where this works alongside teacher marking.

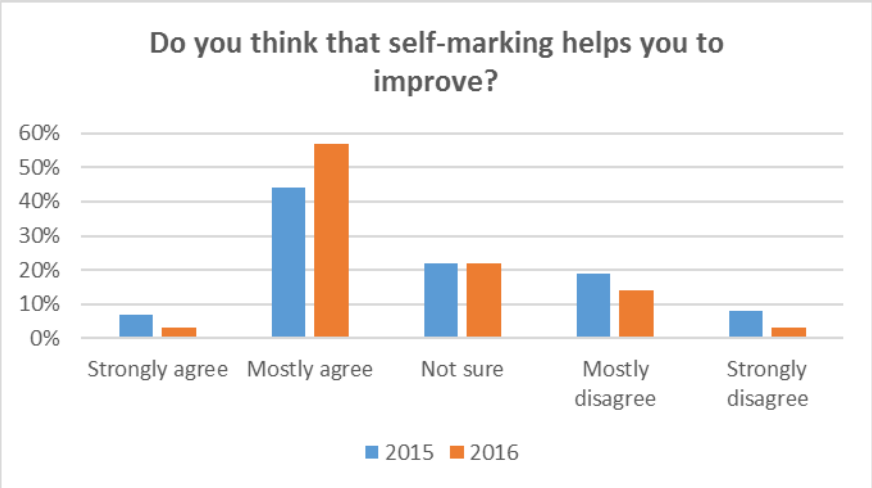


Figure 2

The evidence presented in Figure 2 shows that more students this year agreed that self-marking helped them to improve. This year 60% of students agreed that self-marking helped them to improve (combining data for 'strongly agree' and 'mostly agree'), compared to 51% last year. Last year 27% disagreed that it helped them improve, whilst this year this was reduced to 17%. However, in terms of the extent to which this suggests greater engagement with self-assessment is slightly complicated by the fact that 4% fewer students 'strongly agreed' that self-marking helps them to improve, although in numerical terms this would only indicate a decline of two or three students.

The open responses provide a greater insight into how the students experienced self-assessment, and thus allow for further measures of efficacy. Of the students who 'strongly agreed', two provided justifications, both of which indicated the value of learning from self-assessment. One comment stated, 'It makes you think about how it's marked and how to improve so you remember better for next time', and the other comment stated, 'You can pick up on your mistakes and learn by them'. This to an extent differs from some of the justifications for 'mostly agree' which appeared to use more passive language to describe the

benefits of self-marking, particularly the use of 'see' (to 'show' or 'see' were the most commonly used verbs in these responses). Holding a view that self-assessment helps students to 'see' their mistakes arguably indicates less engagement in the task compared to the student who considers self-marking to help them 'think', 'remember' and 'learn' (to quote the 'strongly agree' students). Indeed, as student responses moved down the Likert scale there appeared to be a changing understanding of the purposes of self-marking being presented. For example, of the 'not sure', 'disagree' and 'strongly disagree' categories, the most common concern cited was of the accuracy of self-assessment (8 out of 23 responses). This might suggest that those who consider self-marking to be useful and perhaps engage with it more think of self-marking as a process to learn from, rather than as an accurate system of measurement. Indeed, this can be seen within the context of a concern highlighted in the literature review about differentiating between the 'useful' purpose of formative assessment compared to the 'accurate' purpose of summative assessment (for example, by Harlen and James, 1997). Thus, to further increase student engagement with self-assessment, it might be important to emphasize the different purposes of formative and summative assessment in terms of utility and reliability.

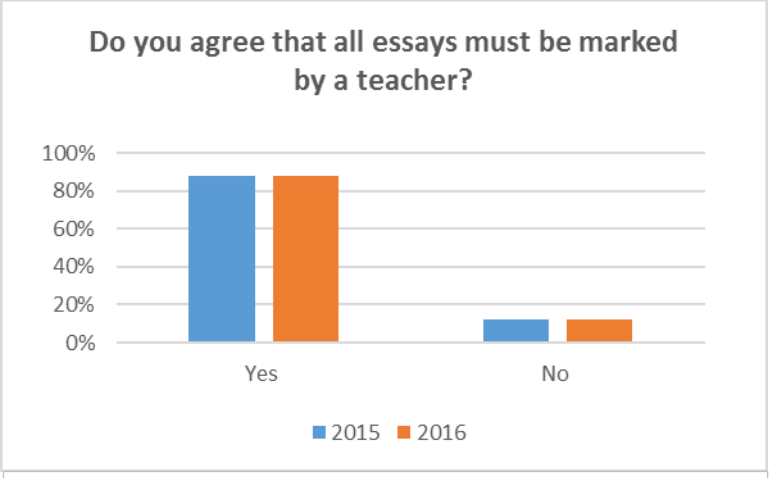


Figure 3

Figure 3 shows that 88% of students in both years agreed that all essays must be marked by a teacher, suggesting no change in the extent to which students have become more independent in the assessment process. Combined with the data presented in Figures 1 and 2, this suggests increased engagement with self-assessment has not led to a consequent "break" in reliance on the teacher which Sadler requires in order for students to become 'self-monitoring' (1989: 120).

However, the open responses revealed a more complex picture behind the quantitative data. For example, it could be assumed that if a student selected 'no' (they did not think all essays must be teacher-marked) then this must be an indicator that they had become self-monitoring. However, this was evidently not the case; for example, two students who had selected 'no' did so because they would feel 'ashamed' for a teacher to see 'bad' work (to paraphrase both responses).

Furthermore, a change in student thinking about assessment could be concealed by the restricted nature of the quantitative data presented here. Combined with the evidence from Figure 1, which showed an increased preference for self-marking, it could be the case that students can show greater engagement with self-marking and might still be reliant on the teacher for confirmation of this. Indeed, the role of the teacher in "confirming" student assessment was the third most commonly cited reason for why teachers must mark all essays (six responses). Thus, the quantitative data might suggest identical levels of teacher reliance, but it might be that students now felt that they have a greater role *alongside* the teacher in the assessment process.

The most commonly cited reason given for teacher marking was for greater accuracy (ten responses), with the second most common being to measure current progress (eight responses). This might suggest that many students see teacher marking in a primarily summative way- in terms of accuracy and grading- such as the comment, 'The teacher can better gage your mark'. Indeed, there is much less mention in this question compared to the previous two of what the students might do with teacher feedback (only three students claimed that teacher feedback helped their learning or understanding), with the comments focusing instead on the quality or nature of the teacher feedback itself rather than how this might impact on their learning process. If students do primarily see teacher assessment as summative, then this might suggest a limited sense of their own responsibility in the assessment process, and that essays are simply completed as a 'finished product' (Price et al., 2010: 280) and handed over for teacher judgement to provide a measure of progress.

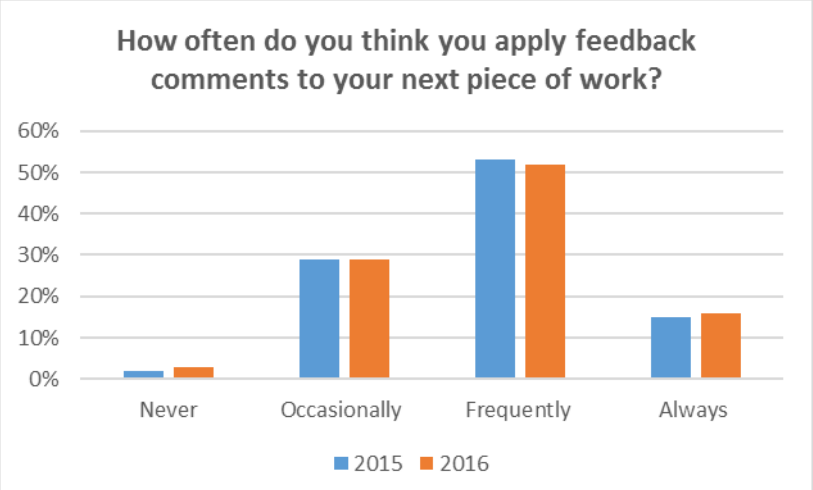


Figure 4

The similar results shown in Figure 4 in response to the regularity in which students reported applying feedback comments seems to be a further indicator that confidence in the efficacy of self-marking has not translated to deeper changes to student feelings of responsibility within

the assessment process. Despite students being encouraged to apply feedback, it did not form a formal part of classroom formative assessment tasks, and an important lesson to learn from this data would be to place a greater emphasis on this next year. Indeed, findings from the second questionnaire showed that most students found target-setting useful, with 41 students identifying it as 'quite useful' or 'very useful' compared to only two students who did not find it useful.

In some ways, it could be argued that this question acts as a control for the changed views on self-marking (although it was not intended to do so). The ways in which the department approached target setting did not substantially differ between the two years, while the approach towards self-assessment did, and so the similarity in responses to this question might provide some control to measure the significance of the changed views on self-assessment, and when seen in this light, weight is leant to the conclusion that there has been an increase in student confidence in self-assessment as a result of a change in department policy, however this is necessarily concluded with caution.

Questionnaire 2

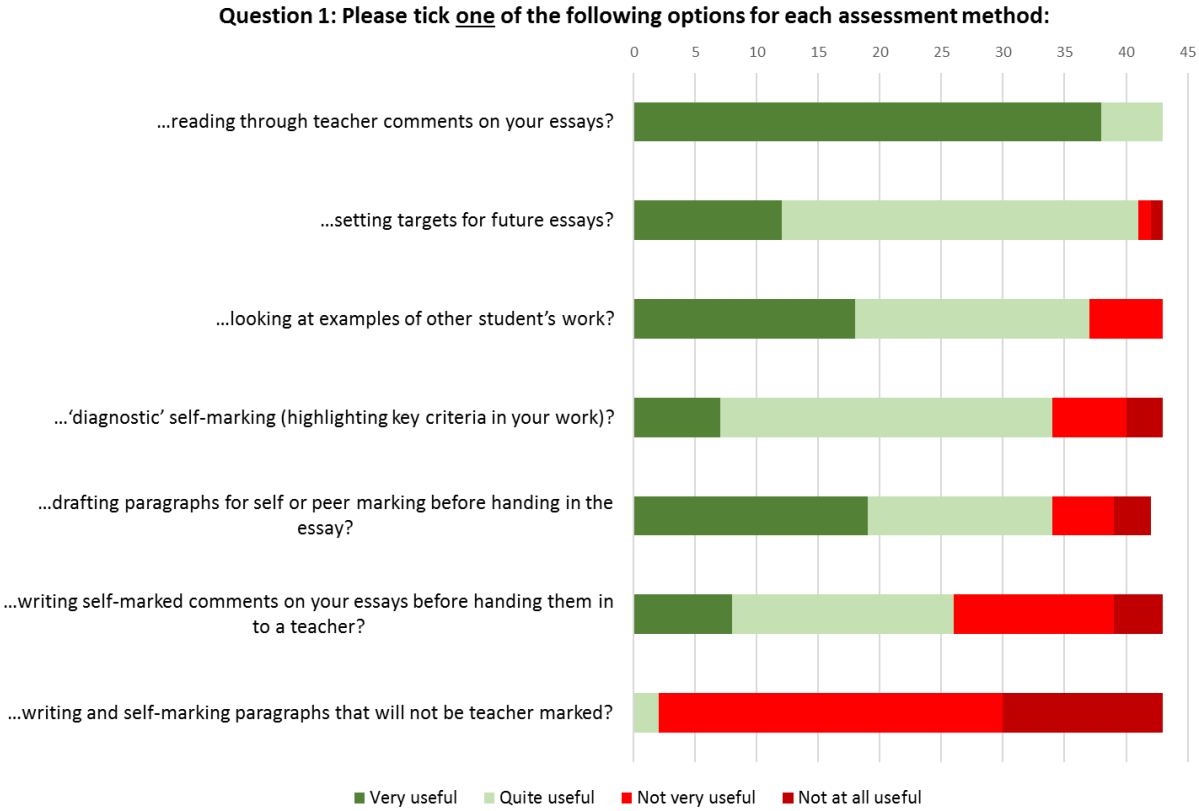


Figure 5: shows the actual number of students who selected each option on the Likert scale (from 'very useful' to 'not at all useful') for each assessment method out of a total sample size of 43. The bars are presented in descending order from the assessment methods students found most useful, to least useful.

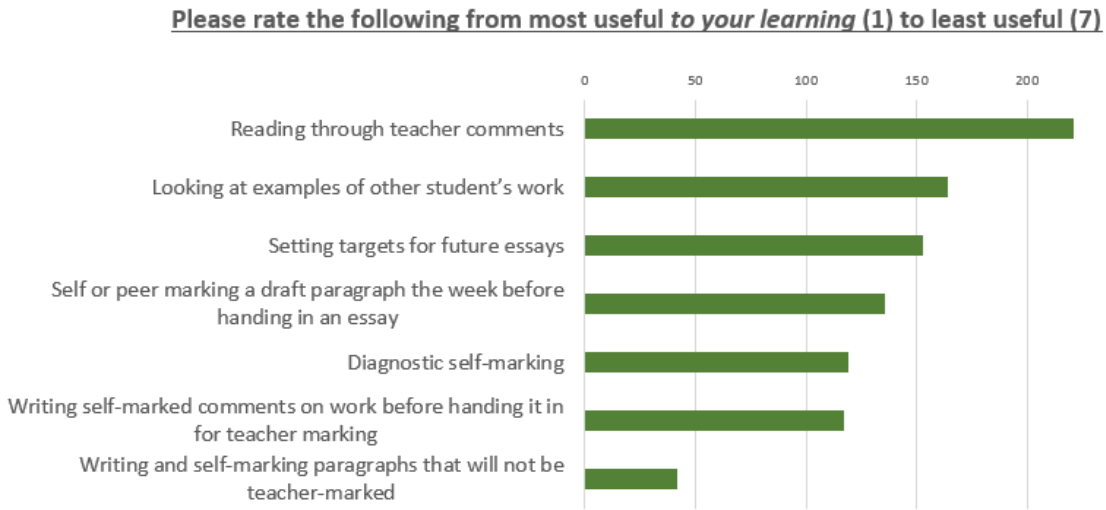


Figure 6: represents students' rank ordered responses by ascribing a descending numerical value from '7' to the top-ranked option, down to '1' for the lowest ranked option. Again, the bars are presented in descending order from the assessment methods students found most useful, to least useful.

In the second questionnaire, students were asked to identify how useful they found specific formative assessment practices (as expressed on a Likert scale), and then to rank order them. A comparison of the two provided internal consistency. This questionnaire was primarily designed to answer the second research question however it also provided additional 'slices of data' (Cohen et al., 2000: 76) to answer the first and third research questions.

Despite most students in the first questionnaire selecting that their preferred method of assessment was teacher and self-marking, the four self-marking options here all appeared at the bottom of both the Likert scale and rank ordering question (as shown in Figures 5 and 6). At first, this may appear to be an inconsistency, but in fact is congruent with the larger picture developing from the data that most students do value self-marking, provided it is carried out alongside teacher marking. For example, despite appearing on the bottom half of the rank order, the only option that was given more negative than positive scores was the option to write and self-mark paragraphs that will not be teacher marked. Drafting paragraphs for self or peer marking before handing in the essay for teacher marking was the fourth most popular option, and yet still achieved 34 positive scores (of 'very useful' or 'quite useful') compared to only eight negatives ('not very useful' and 'not at all useful'). This is consistent with the responses to the first questionnaire, where no students selected 'self-marking only' as their preferred form of feedback.

In terms of the justifications for the most preferred assessment activity -reading through teacher comments- of the 25 written responses to this question, four justified their choice on the grounds that teachers knew more, and were therefore in a better position to mark, such as the comment, ''They know best…' and a further four felt that teachers provided the more

accurate marking. Both of these justifications seem to see teacher marking as a summative activity. The remaining 17 responses all mentioned the role the student played in this process, and in some way argued that teacher comments helped them to then improve their work, as shown in the comment, 'Helps me to improve for next time'. This seems to reinforce the findings from the first questionnaire that students claim a preference for a shared role of the teacher and student in the assessment process. Nonetheless, many of these comments still seem to see teacher feedback as a 'gift' (Price et al., 2010: 280), albeit one which the students then need to do something with in the future. Eight of these claimed that teacher comments provided clear targets that students could then work on, while nine comments did not state specifically how teacher comments helped them improve, other than that they did. The weaker reasoning offered as to how teacher marking helps students improve was also seen in the first questionnaire, which again might suggest that although students like teacher feedback because it is seen as accurate the learning gained from the process might be more limited.

Responses to the three most popular options of teacher-marking, target-setting and model answers all showed evidence of students valuing these for providing clarity, as shown in the comment, 'Gives a clear indication on the areas that need improvement', and indeed a lack of clarity was often cited for the less popular self-assessment activities, exemplified by comments such as 'I don't really know as well as the teacher' and 'Don't know if it is right or not'. Comments from the group interviews reinforced this point; Clara for example said that, 'I find it hard to actually mark my own and see, see things I need to change' and in another interview Emily said that 'you just can't tell whether you've done something right or done something wrong', again suggesting that students consider these activities to be less effective because of the lack of clarity self-assessment offers. However, an application of Edwards'

(2014) quadrant model could provide an insight into this characterization of assessment tasks and a different measure of effectiveness. Edwards' 'third quadrant' of task sequencing is the most challenging, and allows learners to develop higher-order thinking, however she argues that for students to make this much progress, this learning needs to take place in an environment of 'ambiguity and risk' (2014: 23). Thus the validity of these student responses could be questioned when the data is considered in light of Edwards' theory in so much as, although the student may cite a preference for teacher marking and target setting, they may do so because these are relatively low-risk tasks which make them feel more comfortable. Although the question asked students which of these activities were most *useful* to their learning, it is of course possible that they were instead really answering which activities they simply preferred, thus challenging the validity of the responses. Indeed, negative comments surrounding target setting, teacher marking, and looking at model answers were rarely negative because of the perceived difficulty of the task. Of the responses to 'looking at examples of other students' work', only two cited difficulties in applying these lessons, no students cited any difficulty with teacher feedback, and only two students cited difficulties in applying targets to future essays. This compares to five students who stated that they found writing self-marked comments on their own work before teacher marking difficult. This is reinforced by the findings of Stefani (1994) and Orsmond et al. (2000), as both of these studies found students to report finding self-assessment 'challenging' (Stefani, 1994: 73; Orsmond et al., 2000: 29). According to Edwards' theory therefore, a case could be made for the efficacy of these self-marking activities precisely because students find them ambiguous, however this does highlight a tension between taking student views at face value, or interpreting them in light of other research which requires a leap of reasoning which risks distorting student views.

Furthermore, some of these 'less-challenging' formative assessment activities that students cited a preference for could be encouraging negative perceptions of what it means to improve, thus further complicating the extent to which they should be seen as effective. This can be explored through student perceptions of the usefulness of model answers, as in both the questionnaires and group interviews the discussions around model answers particularly provided interesting insights into how students felt they made progress, and thus merits an extended discussion. For example, in the student responses to 'looking at examples of other student work', 7 of the 22 'positive' comments indicated that they liked using model answers because it allowed them to copy or imitate in order to improve, whilst only four students expressed a use of model answers to provide them with new ideas. This concern around the use of formative assessment to learn by imitation is also highlighted by Orsmond and Merry in their study of undergraduate biology students, in fearing that some assessment practices risked students 'not developing into biologists, but merely becoming mimics of biologists' (2011: 133). In one group interview, Amelia stated that model answers are 'perfect essays' and are 'really helpful because you can see what it is supposed to be like'. Laura supported her point, describing 'a perfect essay and exactly how to structure it', and this was reinforced by Clara's statement that, 'I just like to see what would be the best essay to get like an A grade or something and compare with your work to see what you need to change'. The language here suggests a process of using model answers to learn by transmission – that there is one set way of writing an essay and this can be learned by looking at example answers rather than the harder work involved of "discovering" knowledge through more ambiguous self-assessment tasks. Amelia goes on to explain that 'even the starting sentence has to be a certain way' reinforcing Burnham and Brown's observation that some students see essay writing as a process of 'colouring by numbers' (2004: 9).

Other students also described learning from model answers as a process of transmission, but not in a positive way. For example, Timmy said that his 'problem' with model answers was that if he feels his essay 'doesn't align with that of the model answers I've done something wrong or I feel really depressed because I haven't done it to this standard or I haven't structured it in this way'. Estella reinforces his point by saying 'you feel too much pressure to kind of conform to that mark scheme', and Gus also builds on the theme, stating that model answers are 'a bit limited'. Gus sees the purpose of model answers as being to 'copy those in a sense', which he cites as being at odds with the aim to help students become independent learners. Thus it seems whether or not the students find model answers useful, there is nonetheless some consistency in identifying the learning process as being one of imitation, thus suggesting that model answers provide a relatively shallow learning experience, and therefore arguably, a less useful learning process than seems to be indicated by the quantitative data in Figures 5 and 6.

Although some other students placed learning from model answers into a more constructivist context, they were in the minority. Rob for example discussed the benefits of seeing a 'variety of different responses' to learn from, and Estella claimed that the difference between GCSE and A level was that at GCSE students were 'given structures to follow for essays' but at A level 'you can structure it in different ways'. Others saw model answers as a scaffold to help them start the essay, for example, Rosanna stated that 'I often sit there and don't really know how to start it off or even what sort of points to put in or how to arrange it, so it's helpful just to look at a model answer'. This might suggest the essential role a lower-risk formative assessment task has for some students in at least helping them gain a foothold on the task, and to some extent challenges Edwards' view that the most learning happens in the more ambiguous tasks because for some students the pressure of learning in that environment

might be too much (although this is more a question of emphasis than contradiction as Edwards does state that this type of learning is difficult for some students, 2014: 23). Thus in these examples, the use of model answers both offers a scaffold for students struggling with starting an essay, but also an opportunity to analyse different methods of writing for more confident essay-writers. However, it is clear that the transmission model of learning seems to dominate, and perhaps by engaging students in a discussion about the purpose of using model answers this would encourage a more critical use of them.

That said, it should be considered that an activity which might be at first seen as more open-ended and ambiguous and thus leading to higher-order learning might not necessarily do so. For example, of the sixteen students who wrote positive justifications for the use of self-marked comments alongside teacher marking, ten of these could be seen as demonstrating a broadly constructivist interpretation of the tasks, with five saying that it helped them analyze their work in more depth and thus improve, and five also commented on the advantages of working alongside the teacher in developing feedback. However, four of the positive comments instead indicated that they felt self-marking was more about 'proving' something to the teacher, for example the comment, 'Shows that you have knowledge', and the comment, 'Allows me to let teacher know what I feel I've done good/bad'. This is reinforced by some of the responses to the task of drafting paragraphs for self or peer assessment before teacher marking where 6 of the 21 positive comments cited the advantage simply of breaking the essays up and writing it in stages, thus making measures of efficacy difficult to judge from quantitative data alone.

The group interviews showed further differences in how students conceptualized the process of teacher and self-marking. Laura stated that although she does 'find giving myself feedback helpful' it is 'definitely more so with teacher comments' because 'they are going to give me exactly what I need'. The language of 'give' certainly reinforces the view others have made of teacher feedback being seen as a 'gift' (Price et al., 2010: 280), thus suggesting the measure of 'helpful' as being helpful towards producing a product, rather than helpful to learning or understanding, and perhaps therefore primarily seeing the whole feedback process as more summative than formative in approach. Equally, this could be an inaccurate interpretation if she values the accuracy of teacher feedback for providing clear formative guidance. Emily offered a similar view, stating that having both self-marked comments and teacher comments 'gives you more confidence in yourself', and that 'it helps you to know that…your judgement is correct', which again suggests less of a focus on the benefits that self-marking might have as a process compared to an overriding concern with the accuracy of it as an outcome. Bennett argues that the division between formative and summative assessment is 'potentially problematic' (2011: 7) and he concludes that the divide might be an artificial one. However, I think this is unhelpful, and that students might benefit from greater clarity of the differences between the two. For example, if Emily had a clearer understanding that the purpose of self-marking lay in the process rather than the outcome, then she might have fewer concerns with the accuracy of her comments as she would feel that she could learn something through the process of writing them, and then she can see teacher marking as providing a summative judgement. Indeed, arguably the source of much unease about formative assessment is that students view it summatively, hence their then more understandable concern with accuracy. This is to an extent blurred by Sadler, whose ultimate aim is for students to be able to judge the quality of their work, and thus in effect to be able to make a summative judgement of it,

however in the early stages of formative assessment this might not be desirable or helpful to students.

However, some students did demonstrate a more constructivist approach to learning from self and teacher assessment, although again they were a minority. Rob stated that both self and peer feedback were useful to him, claiming that, 'hearing what everyone has to say about work is helpful' suggesting a genuinely social view of learning in the vein of Vygotsky's theories where knowledge is constructed rather than given. However, the fact that relatively few students demonstrated such views on learning might suggest that the assessment policy this year had little influence in developing this, and indeed as a wider concern in interpreting this data, it is perhaps impossible to identify the extent to which attitudes towards learning have been conditioned by specific classroom practices this year, or have developed much earlier in their schooling. For future practice, discussing with students the purposes of using self-assessment might be beneficial as it seems that for some students, formative assessment activities have been transplanted into existing behaviourist theories of the learning process and thus it cannot be assumed that only changing activities and behaviours in the classroom will automatically lead to a change in mindset without expressly questioning these assumptions with students in the first place.

Final consideration is given to the only response to be judged ineffective (in so much as it received more negative than positive scores on the Likert scale)- writing and self-marking paragraphs that will not be teacher marked. Of the 31 negative comments provided for this option, 12 of these were coded for concerns with the lack of accuracy that self-marking in isolation risks. This suggests that not only do most students find this method ineffective, the fact that it is couched in terms of the accuracy of the product (similarly to the previous

questions) rather than the process also suggest that there has been ineffective communication to the students about the purposes of self-assessment, which again returns to concerns of utility and reliability highlighted in the literature review.

Overall it can be concluded from the data provided by the second questionnaire that while students have reported finding most assessment activities effective, when exploring the reasoning behind this judgement, a different picture emerges. Whilst some students do show steps towards independence and self-monitoring, others show a continued reliance on the teacher. It is perhaps unfair to make the ultimate test of efficacy a complete break from teacher support – constructivist classrooms still see a role for the teacher and perceive the learning process as a shared endeavor between teacher and student – however many formative assessment tasks have been interpreted by students in a more 'behaviourist' way of either learning through imitation or teacher instruction which does undermine the overall effectiveness of the policy if a stated aim of formative assessment is to undermine these beliefs. This can perhaps be tackled in the future through more discussion about the purpose of formative assessment tasks to encourage a more critical and thoughtful approach.

Questionnaire 3

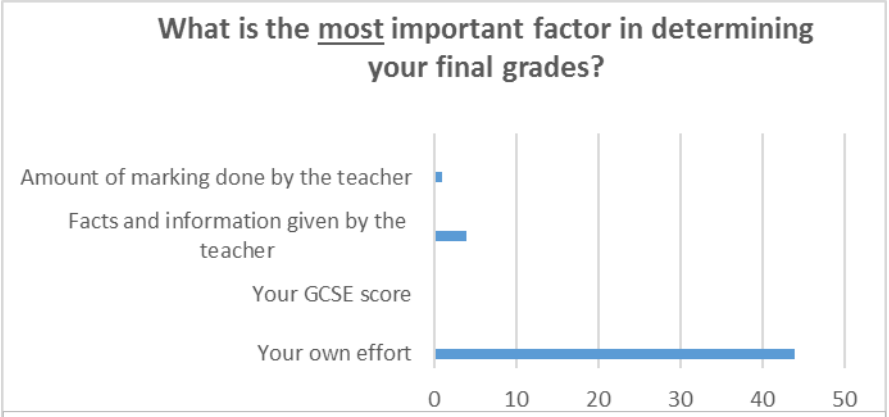


Figure 7 (actual number of student responses shown, out of a total sample size of 52)

Figure 7 clearly demonstrates that most students surveyed felt it was their own effort that was the most important factor in determining their final grade. Only one student selected the 'amount of marking done by the teacher' and did not leave a comment to explain why. Four students chose 'facts and information given by the teacher', and three of these students left comments to justify their choice, all three of which primarily conceptualized the course as fact acquisition, and therefore reasoning that facts must be the key determiner of success. Of the 44 students who selected 'your own effort', most of the responses simply reiterated the importance of effort without providing much justification for it, including comments such as, 'The more personal effort put in the better the grade', and 'If you don't put the effort in, you'll most likely fail'. Of the 16 students who did provide justifications, only three cited the importance of motivation, whilst 13 indicated that effort meant putting in more time or completing extra work, thus perhaps not demonstrating *responsibility* as so much an understanding that the more time put in, the better the outcome is likely to be.

When considered alongside the second questionnaire, further complications arise in concluding that these results must show a significant degree of student responsibility. Despite a clear feeling among the students surveyed that their own effort was the most important factor in determining their final grades, this sits alongside the finding from the same students that their preferred methods of assessment tend to be those that involve the least effort on their part (teacher marking, looking at model answers, setting targets). However, whilst this question asks about the importance of student effort in determining final grades, this does not necessarily answer questions about student feelings of responsibility within the everyday process of classroom assessment. Indeed, effort here could be interpreted as revising and working hard in lessons, and perhaps this model of effort has not extended also to engagement in the assessment process for all students. This might suggest that some students therefore see class-set essays as a teacher's responsibility to assess, which would explain how students can feel ultimately highly responsible for their own grade in the exam, but not so much for classroom-assessment. Discussing with students the different purposes of classroom assessment compared to that of the final exam may help differentiate between the summative purpose of national exams and the formative purpose of classroom assessment, and thus encourage the student to take a more central role in the assessment process. That students clearly do identify the importance of their own effort can then be used to demonstrate how much more learning could take place if they also applied this effort to the assessment process.

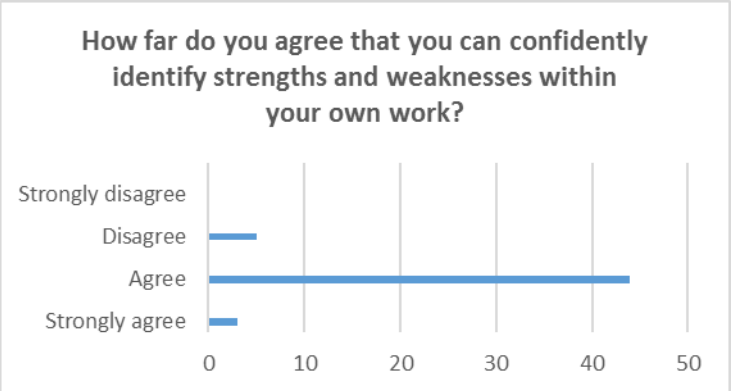


Figure 8

Figure 8 shows that most students (44 out of 52) agreed that they could confidently identify strengths and weaknesses within their own work. This does to some extent challenge findings from the first and second questionnaires which suggested many students were concerned with the accuracy of self-marked work. However, this might be reconciled by considering that students might consider self-marked work to be *less accurate* than teacher assessment, but not necessarily inaccurate per se. An alternative interpretation of this discrepancy could be timing, as students responded to this questionnaire four months after the first questionnaire, lending weight to Dochy et al.'s conclusion that the 'accuracy of self-assessment improves over time' (1999: 337), although this view is less convincing when compared to the data collected from the second questionnaire. That said, of the 38 'agree' comments, 17 of these emphasized that they were more confident in identifying strengths and weaknesses alongside some form of teacher support or guidance. This seems to reinforce earlier findings that students do find self-marking useful, but also value teacher support and guidance alongside it. This further confirms the conclusion that perhaps teacher guidance and self-marking need not be mutually exclusive if teacher guidance throughout the year can help students achieve a self-monitoring standard by the end.

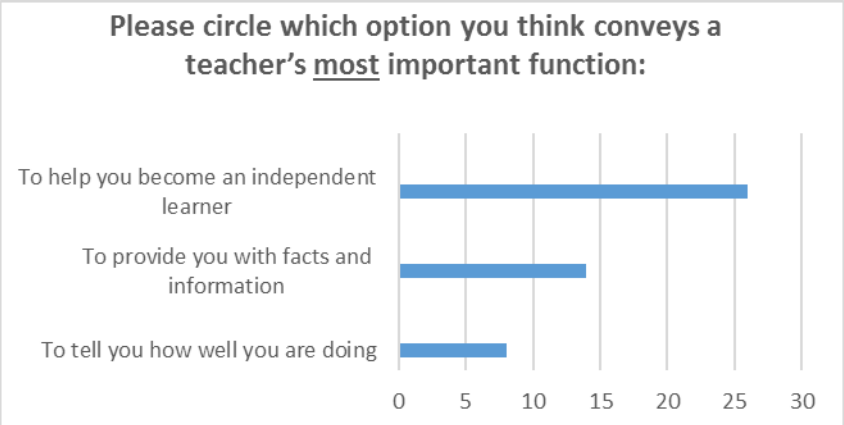


Figure 9

Figure 9 shows that most students (26 of 52) selected that the teacher's most important function was to help students become independent learners, suggesting most students do feel responsible for their own learning. However, 5 out of the 21 responses provided for this option considered being independent to be useful later in life, for example at university, rather than a skill that was also relevant to the A level course, thus limiting the extent to which the quantitative data provides an accurate picture of student feelings of responsibility at their current level.

The students who selected the option of 'to provide you with facts and information' tended to less clearly rationalize why they had selected this. Examples of such responses included, 'They need to teach us stuff', and 'That is their job'. It could be considered that for many students, the importance of providing facts was self-evident, and thus the justifications were often weaker. Indeed, the less clearly rationalized and explained nature of these responses might indicate that this option is the more deeply held assumption among students, and they are therefore less able to rationalize their reasoning as it is held as engrained wisdom. Of the 11 written responses, six of them in some way indicated that this option was the most important

because the course was primarily made up of facts, therefore passing on these facts must be the most important role for teachers; such comments included, 'The facts and information form the basis of much of the work'. Some comments explained a 'transmission' model of teaching, such as 'it is a teacher's primary job (I believe) to pass on this information/facts'. Although these responses seem to indicate a passive role for the student in the learning process (with limited responsibility), the open responses revealed a different picture in some cases. For example, one student who selected 'facts and information' wrote, 'The teacher gives you a basic understanding of a subject so you can research different aspects of that thing in more detail'. Although the teacher still has an important role here for the student in transmitting knowledge, clearly the process for this student does not end here, and instead this is seen as an introduction to a topic for the student to then pick up the responsibility to research it in detail, thus despite the fact they had opted for an option which looks passive, their reasoning was not so, demonstrating the limitations of the conclusions that can be drawn from quantitative data alone. Further comments in response to this question that challenged some of my preconceptions about the meanings behind these options included one response to the option 'to tell you how well you are doing', where one student wrote, 'Becoming an independent learner is important but group exercises and being a 'team player' is also important, so there needs to be a balance of encouraging independence but also encouraging to speak to class mates, challenge one another, and discuss ideas'. I would interpret this as a strongly constructivist view of learning, and yet to have only counted this student's contribution from the quantitative data – where they selected option 'A' (which I would consider to be a less constructivist interpretation of the teacher's role) then this insight would have been lost. This point also highlights a tension in the process of formative assessment more widely between the "contagious minds" of Vygotskian constructivist learning where students improve based on co-operative learning compared to Sadler's more

independently-minded self-monitoring student, and returns to the issue of the *extent* to which students must be independent and responsible within the assessment process.

Overall, the data from this question reinforces the findings from the first question that although the majority of students cite the importance of independence and responsibility, it is less clear how far this then extends to their role in the assessment process.

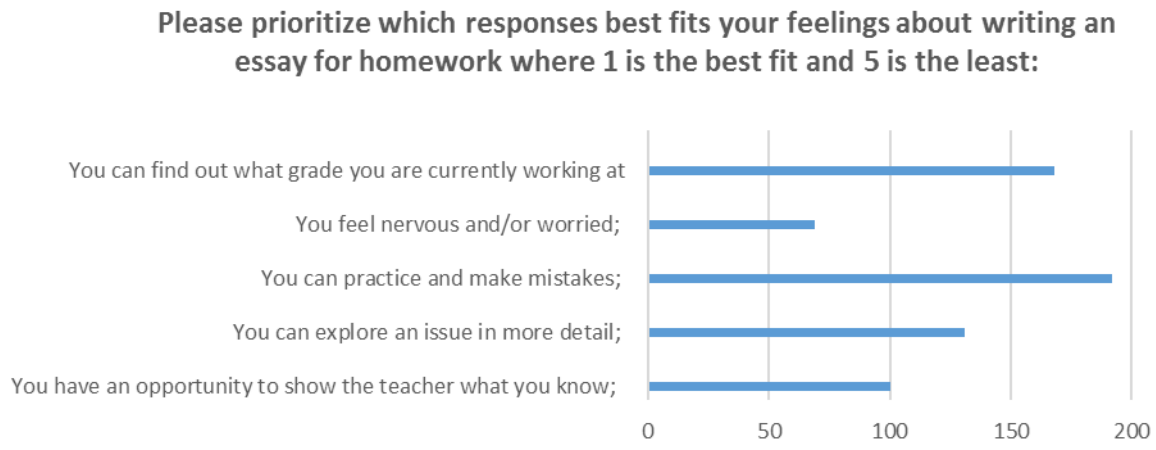


Figure 10: represents students' rank ordered responses by ascribing a descending numerical value from '5' to the top-ranked option, down to '1' for the lowest ranked option (as in Figure 6).

Figure 10 shows a rank ordering of student responses to writing essays. Most students (27 out of 44) selected 'you can practice and make mistakes' for their first preference, and 12 of 32 written comments discussed the importance of making mistakes in the learning process, such as '…mistakes are useful and can learn from them…'; 'If I make a mistake I can work out how to fix it'; and, 'I feel that making mistakes allows you to improve the quickest'. Although this data might suggest that most students see essay writing in class as part of a continuous learning process rather than seeing each essay as a single final and perfectible product, those

students who do not see essays in this way make up an important minority. For example, although far fewer students stated that essays primarily made them feel 'nervous and/or worried', their comments stood out for the extent of the emotional response they felt to essay writing. Such comments included, 'I do feel very nervous and worried as I love history and want to do well but am nervous I'm not good enough' and, 'Always worried that you have failed'. The students that expressed emotional concerns all linked these to achieving a standard, whether that is 'good enough' in the first comment, or 'failed' in the second, and other comments specifically referenced achieving certain grades. This would suggest that the source of this anxiety might primarily be the grade, or at least some "measurable" part of the feedback, and this could be corroborated by theories of ego-involved motivation, as explored in the literature review. Tackling these feelings might to some extent be achieved by continually characterizing essays as part of a larger learning process and would again merit further discussion about the purposes of formative and summative assessment (as also considered in the literature review). It seems that grade focus is still an important preoccupation – it was the second most important factor selected here. This preoccupation with grades and measurement might indicate a view of assessment as the teacher's responsibility to provide a grade rather than of student responsibility to improve. Although the most commonly cited reason here was to 'practice and make mistakes', this does not guarantee student responsibility if students think these mistakes must be pointed out by the teacher in order for them to improve, which would be congruent with the findings from the other questionnaires that seemed to place much more confidence in teacher-set targets than student-set targets.

Student suggestions

An additional way of exploring the extent to which students feel responsible within the assessment process might be to see how far they can critique it and offer improvements. Exploring student suggestions for how to improve the policy not only offers some insight into the extent to which they feel responsible for it, but the suggestions themselves offer further insights into how students view their role in the assessment process.

Questionnaire questions that asked students to suggest improvements to the current assessment policy were left blank by 24 students, and many others wrote that they were happy with the current policy. At face value, this would indicate satisfaction with the current policy. However, when considered in light of some of the issues raised in the literature review regarding student power in the classroom, it could indicate that students found it difficult to offer suggestions or alternatives because they were not used to thinking critically about classroom assessment if they saw it as something which is decided by the teacher. Students in the group interviews however were more forthcoming, and the issue could therefore be one of data collection, as the discursive nature of the group interview was more conducive to exploring alternative possibilities as students could build on each other's ideas compared to the private and individual completion of a questionnaire. Suggested improvements from both the questionnaires and group interviews included: writing targets on the top of essays, keeping a written record of feedback on a separate list (a "progress tracker") and the use of one to ones to discuss progress with teachers. The improvements suggested provided further insight into how the students conceptualized feedback and improvement. For example, in one group interview Clara suggested keeping 'a sheet of paper where we always write down the things we need to improve on so then each time we can go back…and then tick it off'. And in

another group interview Estella suggested a similar measure. Although this suggests a degree of responsibility being taken in so much as they are both suggesting a new direction for the department's practice, the improvement itself suggests a relatively narrow interpretation of improvement whereby there is a linear list of skills which can simply be acquired and ticked off, a concern also highlighted in Nicol and Macfarlane-Dick's study (2006: 209).

Indeed, this view of formative assessment as accumulating a 'list' of targets was also mentioned by Rosanna in a group interview. She said that she has key targets as a 'list in my head but I'm not sure how to approach it at all', thus suggesting that although she can use some of the formative assessment tasks to identify improvements, these tasks in their current format do little to help her then apply these to future work. Evidence of this problem in "closing" the feedback loop was also identified by Beaumont et al. (2011) and Orsmond and Merry (2011), who also found that students struggled to apply feedback to future work. This might suggest that to make student ownership of assessment more coherent, students could use directed time in class to practice applying these targets, otherwise their engagement in the assessment process remains stuck at the "diagnosis" stage, with the teacher simply assuming that these "diagnosed" targets will be applied to future work. The use of directed time was also suggested by Rob in one of the group interviews, where he suggested writing another paragraph shortly after receiving feedback in order to apply targets nearer to the event, without having to wait for the next essay to be set.

Estella suggested that more one to ones would improve the department's formative assessment policy. She highlighted the importance of being able 'to talk through essays with your teacher'. The other students in her group were also very supportive of this suggestion.

Rosanna for example pointed out the key problem with using self-marked comments and then teacher comments on essays to establish a dialogue is that 'you can't really, like, keep sort of sending a letter back to keep asking questions' compared to a one to one where 'you could have a full discussion about it so you can fully understand'. Indeed, it seems obvious that the best way to establish a dialogue with students is to talk to them, and yet many formative assessment approaches seem to overlook this approach in favour of written processes such as dialogic marking, which might "prove" a dialogue to an outsider observer, but might be an inefficient way of achieving something that a one-to-one could do more effectively. Price et al. also found students had a 'hunger for more opportunities to have a dialogue with staff' (2010: 284) and Blair and McGinty found that students valued time spent discussing their work verbally with teachers (2013: 470).

However, the use of one-to-ones may not be as empowering as many researchers suggest. Although many studies -particularly those involving higher education students - advocate the importance of dialogue between students and teachers in order to increase the profile of the student within the assessment process (e.g. Nicol, 2010; Price et al., 2010; Carless et al., 2011; Blair and McGinty, 2013; Orsmond et al., 2013); they do not particularly address the context of power relations in which this dialogue might be placed, and in a view by Black previously cited, in that dialogue, the 'teachers' power easily overwhelms' (1993: 81). Furthermore, it might be considered that increased dialogue is no guarantee of increased student independence if it is simply seen as another form of mediated teacher-transmission of marking criteria, and indeed an increase in dialogue could risk increasing the reliance on the teacher, compared to advocating greater involvement of students in their own self-assessment. Nonetheless, opening up a dialogue does at least provide a platform for students to negotiate assessment with their teacher, but it is presumably no panacea.

Summary of findings

An overall summary of findings is now provided in response to the specific research questions:

1. **Do students report more engagement with self-assessment compared to the students that I taught last year?**

The quantitative data from the first questionnaire does indicate greater levels of student engagement with self-assessment compared to the students that I taught last year; or perhaps more accurately, students at least report finding self-assessment more useful to their learning compared to the students questioned last year. As has already been discussed, it is difficult to arrive at conclusions regarding the extent of student engagement in tasks when it is unclear what success criteria have been applied when students complete the questionnaires, but at face value at least there was an increased level of student engagement reported, suggesting that our assessment policy this year might be moving in the right direction.

2. **To what extent do students consider current self-assessment strategies to be effective?**

Generally, students do seem to consider self-assessment strategies to be effective. The data gathered from the second questionnaire indicates that the only practice that is currently deemed to be ineffective by students is to self-mark paragraphs that will not be teacher marked, however within the context of the discussion around Edwards' 'third quadrant' of learning in an environment of ambiguity and risk (2014: 23) even this conclusion could be challenged, depending again on the success criteria students have employed in order to judge effectiveness. Again however, the issue of motivation is returned to, and it might be possible to conclude that if students perceive the activity in such a negative light, they will bring

limited motivation to the task, thus undermining efficacy. As has already been explored in both the findings and literature review, discussing with students the differences between the 'utility' of formative assessment compared to the 'reliability' of summative assessment may encourage students to approach some of these more ambiguous formative assessment tasks differently.

Furthermore, despite students reporting that they consider assessment strategies such as reading through teacher comments and reading model answers to be effective, an examination of their open responses and interview data suggest that often a relatively shallow degree of learning is taking place during these activities, thus undermining the extent to which these activities can be considered effective, and again returning to questions of the success criteria employed by students in responding to the questionnaire. In the context of the literature review, it seems these activities are not taking place within a constructivist context for all students, and this merits further investigation.

3. To what extent do students report feelings of responsibility within the assessment process?

All three questionnaires offered insights into this more complex question. Although the quantitative data analysed from the first questionnaire demonstrated increased student responsibility, many of the open responses indicated a continued reliance on the teacher. This seems to be reinforced by the findings from the second questionnaire where more passive methods of feedback such as teacher marking and model answers were more popular than self-assessment. The second questionnaire also revealed behaviourist and imitative interpretations of learning from formative assessment, suggesting a less individually responsible and constructivist approach to learning than one of transmission or learning from

teacher-provided feedback, although this was not a universally held view. The third questionnaire suggested to a fair extent that students did consider themselves responsible for their own learning, but when combined with the findings of the second questionnaire it is less clear how far they see themselves as responsible within the process of classroom assessment specifically, compared to their learning more widely.

Perhaps underlying any judgement to this question ought to be an idea of what an acceptable level of student responsibility within the assessment process ought to be. As considered in the literature review, there is a tension between seeing formative assessment as primarily a teacher's responsibility and a part of pedagogy (e.g. William, 2011) and of seeing it as the responsibility of the student, as could be indicated by Sadler's desire for students to become 'self-monitoring' (1989: 120). Even if formative assessment is seen primarily as student-driven, there is a further tension between the Vygotskian outlook of knowledge constructed through shared interactions, and the more personally reflective model of Sadler's self-monitoring student. Even Sadler's self-monitoring student is primarily seen as an 'end goal' rather than something students are able to do during the learning process itself, further complicating any understanding of how independent and responsible a student ought to be during the learning process. Overall, perhaps the best conclusion that can be drawn in response to this third research question is that most students do consider themselves to have a role in the assessment process, but claim to see this as existing alongside teacher support, thus making a sense of student 'responsibility' perhaps overstated.

Conclusions

To return to the 'concentric spheres of influence' considered in the rationale section of this project, it is perhaps worth adopting Claxton's 'split screen' (2007: 122) vision once again to finally contextualize the findings of this project within the wider national picture, before "zooming in" to consider the next steps to be taken to further develop and improve our department's formative assessment policy.

At a national level, it seems that the central role that assessment has within our education system shows no signs of waning. On the 19th July 2016 Amanda Spielman was confirmed as the next Chief Inspector of Ofsted. It was pointed out by MPs sitting on the education select committee than Spielman had no teaching background, and her primary experience of the education system was through her chairmanship of exams regulator Ofqual[4]. In Morgan's words (former Secretary of State for Education), she chose Spielman for her 'evidence-based approach' and 'system-level thinking'[5]. Her appointment is perhaps indicative of a growing trend for assessment to become dominant over learning in the British educational system, and arguably this reaches our classrooms as a growing trend for assessment *as* learning, rather than assessment *for* learning. Indeed, this could be seen as part of a wider societal 'evidence-based approach' where it seems the only experiences of value are those that are accountable and measurable; whether that is the 'bucket-list' phenomenon, or the National Trust's '50 things to do before you're 11 ¾'[6], the underlying message is that the only activities worth doing are those that are quantifiable, and this of course extends to our education system – the "real" learning is the learning to the exam. Classroom formative assessment sits

[4] http://www.bbc.co.uk/news/education-36723828 [Accessed 7 July 2016]
[5] http://www.bbc.co.uk/news/education-36723828 [Accessed 7 July 2016]
[6] http://www.telegraph.co.uk/news/earth/earthnews/10029981/National-Trust-list-of-50-things-to-do-before-you-are-11--and-where-to-do-them.html [Accessed 7 July 2016]

uncomfortably within this growing trend. Indeed, Black and Wiliam's feedback loop model arguably sits as much within this evidence-based approach as school league tables – a systems theory driven by data which risks causing the "assessment tail" to "wag the learning dog". However, formative assessment can be conceptualized in two ways. In one guise, formative assessment is assessment as learning, where exam skills are transmitted from teacher to student, and the only valuable form of learning is assessment preparation. An alternative way of looking at formative assessment is within a democratic and constructivist context, where it empowers students to take responsibility and control over their own education and learning. If assessment must be such a dominant part of learning (as seems to be indicated by the national context), teachers can have a role in mediating this, and applying assessment within their classrooms in a way which can be positive in encouraging student feelings of responsibility and agency, and to encourage a view of assessment that can be empowering and motivating.

The next steps to take following this project will seek to move our department ever closer to this ideal. The data from this research project has suggested that generally students have found formative assessment useful; however, when exploring why students consider these practices to be useful, a more complex picture emerges, and it is clear that there is still scope for students to take greater responsibility within the assessment process. Whilst some students appear to conceptualize formative assessment in a self-responsible constructivist manner, where they work alongside their peers and teachers to build a more concrete understanding of essay writing, others appear to see formative assessment as a passive process of teacher transmission. It might be safe to conclude therefore that simply changing activities and behaviours in a classroom does not automatically lead to a change in deeper-held values and understandings of the student's role in the classroom and the wider learning

process, and perhaps a more fundamental shift in student responsibility is necessary. The next steps that will be necessary in order to achieve this in practice are now outlined below.

Next steps

The overall aim of this project was to use student views to develop and improve our department's formative assessment policy. Although initially this was to consist of using student views to adapt the current policy, the process of exploring student views on a range of formative assessment practices has revealed that the tasks themselves do not guarantee a certain mindset or attitude, as student interpretations of the tasks vary considerably; indeed, formative assessment practices appear to have been negotiated and adopted by many students within a framework of teacher-dominance and student passivity. Therefore, mediating the detail of assessment policies can presumably only go so far in encouraging students to take greater responsibility in the assessment process. Instead, an attempt to fundamentally shift this balance of power will be made by giving students greater responsibility in setting assessment activities themselves. Involving the students in a one-off event of policy-writing would be one solution, but it may be more effective still to regularly engage students in discussions about the effectiveness of formative assessment tasks, with a view to them ultimately setting formative assessment activities themselves. The process of deciding formative assessment tasks not only transfers power over to the students in order to direct assessment policy, it also engages students in a critical discussion about how they learn from formative assessment tasks, thus in theory increases their engagement with formative assessment and further develops an overall sense of student responsibility in the assessment process.

The more challenging task however is to work out *how* to make this work in practice. The unwillingness of students to offer many suggestions to change the assessment policy in the second questionnaire suggests that handing over responsibility to students to set tasks might be difficult at first if the students are not used to being placed in this role. It might be sensible for the teacher to set some formative assessment tasks for the first pieces of written work and to then discuss with the students how useful they found those activities. This regular reviewing of formative assessment tasks might then provide some basis for students to think critically about the assessment process and to develop the language necessary to discuss it, then ultimately paving the way for students to decide as a class when a piece of work is set how they would like to formatively assess it. It would be hoped that if this could become routine towards the end of the first year, then it will be an embedded practice when students enter their second year.

It has become clear from this project that a policy is much more than a set of practices, it is also a set of principles. It is hoped that by encouraging more student engagement in evaluating classroom practice, my teaching practice will move closer to the fundamental ideal of developing a democratic classroom where every student is encouraged to participate and take responsibility for their own learning and progress.

References

Beaumont, C., O'Doherty, M. with Shannon, L. (2011) Reconceptualising assessment feedback: a key to improving student learning? *Studies in Higher Education*, 36:6, 671-687

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

Black, P. (1993) Formative and Summative Assessment by Teachers, *Studies in Science Education*, 21:1, 49-97

Black, P. and Wiliam, D. (1998) Assessment and Classroom Learning, *Assessment in Education: Principles, Policy & Practice*, 5:1, 7-74

Black, P., McCormick, R., James, M. with Pedder, D. (2006) Learning How to Learn and Assessment for Learning: a theoretical inquiry, *Research Papers in Education,* 21:2, 119-132

Black, P. and Wiliam, D., (2012) Assessment for Learning in the Classroom in J. Gardner (ed.) *Assessment and Learning* (2nd edition), London: Sage, pp. 11-32

Blair, A. and McGinty, S. (2013) Feedback-dialogues: exploring the student perspective, *Assessment & Evaluation in Higher Education*, 38:4, 466-476

Boud, D. and Molloy, E. (2013) Rethinking models of feedback for learning: the challenge of design, *Assessment & Evaluation in Higher Education*, 38:6, 698-712

Burnham, S. and Brown, G. (2004) Assessment without Level Descriptions, *Teaching History*, 115, 5-15

Carless, D., Salter, D., Yang, M. with Lam, J. (2011) Developing sustainable feedback practices, *Studies in Higher Education*, 36:4, 395-407

Clarke, S. (2005) *Formative assessment in the secondary classroom,* London: Hodder

Clarke, S. (2008) *Active learning through formative assessment,* London: Hodder

Claxton, G. (2007) Expanding Young People's Capacity to Learn, *British Journal of Educational Studies*, 55:2, 115-134

Claxton, G. (2009) Cultivating positive learning dispositions in H. Daniels, H. Lauder, J. Porter, with S. Hartshorn (eds.) *Educational Theories, Cultures and Learning: a critical perspective,* London: RoutledgeFalmer, pp. 177- 187

Cohen, L., Manion, L. with Morrison, K. (2000) *Research Methods in Education (*5th edition), London: RoutledgeFalmer.

Dann, R. (2002) *Promoting Assessment as Learning, Improving the Learning Process,* London: RoutledgeFalmer

Dochy, F., Segers, M. with Sluijsmans, D. (1999) The use of self-, peer and co-assessment in higher education: A review, *Studies in Higher Education*, 24:3, 331-350

Ecclestone, K. (2002) *Learning autonomy in post-16 education: the politics and practice of formative assessment,* London: RoutledgeFalmer

Edwards, A. (2014) Designing tasks which engage learners with knowledge, in I. Thompson (ed.) *Designing Tasks in Secondary Education: Enhancing Subject Understanding and Student Engagement,* London: Routledge, pp. 13-27

Galloway D., Rogers, C., Armstrong, D., Leo, E. with Jackson, C. (2004) Ways of understanding motivation, in H. Daniels and A. Edwards (eds.) *The RoutledgeFalmer Reader in Psychology of Education,* London: RoutledgeFalmer, pp. 89-105

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

Hargreaves, E. (2005) Assessment for learning? Thinking outside the (black) box, *Cambridge Journal of Education*, 35:2, 213-224

Harlen, W., and James, M. (1997) Assessment and Learning: differences and relationships between formative and summative assessment*, Assessment in Education: Principles, Policy & Practice*, 4:3, 365-379

Harlen, W. (2012) On the relationship between assessment for formative and summative purposes in J. Gardner (ed.) *Assessment and Learning* (2nd edition), London: Sage, pp. 87-102

Hattie, J. (2012) *Visible learning for teachers, maximizing impact on learning,* Abingdon: Routledge

Hayward, L. (2012) Assessment and learning: the learner's perspective in J. Gardner (ed.) *Assessment and Learning* (2nd edition), London: Sage, pp. 125-139

James, M. (2008) Assessment and Learning in S. Swaffield (ed.) *Unlocking Assessment,* London: Routledge, pp. 20-35

Leahy, and Wiliam, D. (2012) From teachers to schools: scaling up professional development for formative assessment in J. Gardner (ed.) *Assessment and Learning* (2nd edition), London: Sage, pp. 49-68

Nicol, D. and Macfarlane-Dick, D. (2006) Formative assessment and self-regulated learning: a model and seven principles of good feedback practice, *Studies in Higher Education*, 31:2, 199-218

Nicol, D. (2010) From monologue to dialogue: improving written feedback processes in mass higher education, *Assessment & Evaluation in Higher Education*, 35:5, 501-517

Nuttall, D. (2013) Possible futures: using frameworks of knowledge to help Year 9 connect past, present and future, *Teaching History,* 151, 33-44

Orsmond, P., Merry, S. with Reiling, K. (2000) The Use of Student Derived Marking Criteria in Peer and Self-assessment, *Assessment & Evaluation in Higher Education*, 25:1, 23-38

Orsmond, P. and Merry, S. (2011) Feedback alignment: effective and ineffective links between tutors' and students' understanding of coursework feedback, *Assessment & Evaluation in Higher Education*, 36:2, 125-136

Orsmond, P., Maw, S., Park, J., Gomez, S. with Crook, A. (2013) Moving feedback forward: theory to practice, *Assessment & Evaluation in Higher Education*, 38:2, 240-252

Price, M., Handley, K., Millar, J. with O'Donovan, B. (2010) Feedback: all that effort, but what is the effect*? Assessment & Evaluation in Higher Education*, 35:3, 277-289

Sadler, D. (1989) Formative assessment and the design of instructional systems, *Instructional Science,* 18, 119-144.

Skerrett, A. (2010) There's going to be community. There's going to be knowledge: Designs for learning in a standardized age, *Teaching and Teacher Education*, 26:3, 648-655

Stefani, L. (1994) Peer, self and tutor assessment: Relative reliabilities, *Studies in Higher Education*, 19:1, 69-75

Stobart, G. (2012) Validity in formative assessment in J. Gardner (ed.) *Assessment and Learning* (2nd edition), London: Sage, pp. 233-242

Swaffield, S. (2011) Getting to the heart of authentic Assessment for Learning, *Assessment in Education: Principles, Policy & Practic*e, 18:4, 433-449

Taras, M. (2009) Summative assessment: the missing link for formative assessment, *Journal of Further and Higher Education,* 33:1, 57-69

Torrance, H. and Pryor, J. (1998) *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom,* Buckingham: Open University Press

Torrance, H. (2007) Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning, *Assessment in Education: Principles, Policy & Practice*, 14:3, 281-94

Wiliam, Dylan (2011) *Embedded formative assessment*, Bloomington: Solution Tree Press

Internet sources

Bureaucratic marking policies add nothing to pupils' progress (2016) [Online] NASUWT. Available from: <http://www.nasuwt.org.uk/Whatsnew/NASUWTNews/PressReleases/NASUWT_015486>. [Accessed 3 April 2016]

Eliminating unnecessary workload around marking, Report of the Independent Teacher Workload Review Group (2016) [Online]. Available from: <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/511256/Eliminating-unnecessary-workload-around-marking.pdf>. [Accessed 3 April 2016]

Gray, L. (2013) *National Trust list of 50 things to do before you are 11 ¾ and where to do them* [Online] The Telegraph. Available from: <http://www.telegraph.co.uk/news/earth/earthnews/10029981/National-Trust-list-of-50-things-to-do-before-you-are-11--and-where-to-do-them.html>. [Accessed 7 July 2016]

Rhodes, A. (2016) *Young people – if you're so upset by the outcome of the EU referendum, then why didn't you get out and vote?* [Online] The Independent. Available from: <http://www.independent.co.uk/voices/eu-referendum-brexit-young-people-upset-by-the-outcome-of-the-eu-referendum-why-didnt-you-vote-a7105396.html>. [Accessed 3 July 2016]

Sellgren, K. (2016) *Ofsted's next head, Amanda Spielman, rejected by MPs* [Online] BBC. Available from: <http://www.bbc.co.uk/news/education-36723828>. [Accessed 7 July 2016]

Appendices

Content	**Page number**
Head teacher approval letter	78-9
Questionnaire 1	80
Questionnaire 2	81-83
Questionnaire 3	84-5
Example of coded data from questionnaires	86
List of indicative questions for group interviews	87
Extract from transcribed interview	88
Indicative extract from transcribed interviews with analytical commentary	89
Research poster (used at MLT conference and with college TLC)	90

Head teacher approval letter

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 Ernesto Macaro

6-3-16

Ms. Sally Bromley
Collyers College,
Hurst Road,
Horsham,
RH12 2EJ

Dear *Ms. Sally Bromley,*

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 Therese Hopfenbeck. In my final research project, '*Implementing a formative assessment policy in a sixth form history classroom*', I am exploring the impact of consistent self and peer marking on student understanding of assessment and their role within the assessment process.

I hope to collect data on student views of formative assessment between March and June 2016. The data collection will consist of three questionnaires of between 5-8 questions given to students in class. These will be completed anonymously and with the option for students to not participate. Students will also be invited to group interviews to discuss their experience of formative assessment in more depth. These interviews will be audio-recorded, and students will be given a question list in advance so that they can give informed consent to their participation.

Oxford University has strict ethical procedures on conducting ethical research with teachers and young people, consistent with current British Educational Research Association guidelines. As practitioner research however, the University recognises that schools have the highest ethical standards in any event. Therefore only your consent is necessary, and not that of parents. 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, will be made anonymous in all research reports. The data collected will be kept strictly confidential, available only to my supervisor and myself, and not used other than specified without further consent. All audio recordings will be destroyed at the end of the research period, and kept in locked conditions until then.

By participating in the research, the college would be contributing to a project that will deepen the department's understanding of formative assessment, and so contribute towards developing ways of using formative assessment for students in the future.

If you need more information about what is involved, please contact me.

I look forward to hearing from you.

Yours sincerely,

Emma Hawkins

Implementing a formative assessment policy in a sixth form history classroom

Emma Hawkins
University of Oxford, Department of Education

College of Richard Collyer

Hurst Road,
Horsham,
West Sussex,
RH12 2EJ

Ms. Sally Bromley

- ☐ 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.

SB-ABromley
09/03/16

If you would like further information, please contact me, or my supervisor. If you require clarification of the ethical approval, process please contact Chair of Department of Education Research Ethics Committee, Dr Nigel Fancourt.

Supervisor email: *therese.hopfenbeck@education.ox.ac.uk*

Chair of Department of Education Research Ethics Committee, Dr Nigel Fancourt email: nigel.fancourt@education.ox.ac.uk

Please return this form to me.

Thank you for your help.

Emma Hawkins

Questionnaire 1

Please read through the following questions and circle the response which best fits your view. There are no right or wrong answers – I am interested in your honest opinion in response to these questions.

1. **Which type of feedback do you find most useful in order to improve?**
   a) Teacher marking only
   b) Self-marking only
   c) Peer-marking only
   d) Teacher and self-marking
   e) Teacher and peer-marking
   f) Self and peer-marking
   g) A combination of all three

Please explain your answer:

2. **Do you think that self-marking helps you to improve?**
   a) Strongly agree
   b) Mostly agree
   c) Not sure
   d) Mostly disagree
   e) Strongly disagree

Please explain your answer:

3. **Do you agree that all essays must be marked by a teacher?**
   a) Yes
   b) No

Please explain your answer:

4. **How often do you think you apply feedback comments to your next piece of work?**
   a) Never
   b) Occasionally
   c) Frequently
   d) Always

Please explain your answer:

Please tick here if you consent for the data you have provided in this questionnaire to be used in my final essay. All student and staff names and the college name will be anonymized. ☐

Questionnaire 2

I have written this questionnaire to find out how useful students find different methods of assessment. For each question, the criteria I would like you to consider is 'how useful to your learning' you find each approach. Don't worry if you find lots of them 'very useful' or 'not at all useful' – you have the opportunity to prioritize them at the end of the questionnaire.

1. **How useful to your learning do you find 'diagnostic' self-marking (i.e. going through your essays or paragraphs and highlighting key criteria such as 'evidence' and 'analysis')?**
   a) Very useful
   b) Quite useful
   c) Not very useful
   d) Not at all useful

Please explain your decision:

2. **How useful to your learning do you find writing self-marked comments on your essays before handing them in to a teacher?**
   a) Very useful
   b) Quite useful
   c) Not very useful
   d) Not at all useful

Please explain your decision:

3. **How useful to your learning do you find drafting paragraphs for self or peer marking before handing in the essay the following week?**
   a) Very useful
   b) Quite useful
   c) Not very useful
   d) Not at all useful

Please explain your decision:

4. **How useful to your learning do you find looking at examples of other student's work?**
    a) Very useful
    b) Quite useful
    c) Not very useful
    d) Not at all useful

Please explain your decision:

5. **How useful to your learning do you find writing and self-marking paragraphs that will not be teacher marked?**
    a) Very useful
    b) Quite useful
    c) Not very useful
    d) Not at all useful

Please explain your decision:

6. **How useful to your learning do you find reading through teacher comments on your essays?**
    a) Very useful
    b) Quite useful
    c) Not very useful
    d) Not at all useful

Please explain your decision:

7. **How useful to your learning do you find setting targets for future essays?**
    a) Very useful
    b) Quite useful
    c) Not very useful
    d) Not at all useful

Please explain your decision:

8. **Please rate the following from most useful *to your learning* (1) to least useful (7)**

Assessment method:	**Number**
Diagnostic self-marking	
Writing self-marked comments on work before handing it in for teacher marking	
Self or peer marking a draft paragraph the week before handing in an essay	
Looking at examples of other student's work	
Writing and self-marking paragraphs that will not be teacher-marked	
Reading through teacher comments	
Setting targets for future essays	

Please explain your decision:

9. **Are there any other assessment methods which you have found useful or not useful?**

10. **What changes would you make to the History department's assessment policy next year?**

Please tick here if you consent for the data you have provided in this questionnaire to be used in my final essay. All student and staff names and the college name will be anonymized.

Questionnaire 3

This questionnaire is designed to understand your views on the following topics. There are no right or wrong answers, so please answer as honestly as possible. There is no need to put your name on this questionnaire, however if you would like to discuss any of the topics raised here then please do speak to me afterwards. The data will be used both by the History department and in my dissertation.

1. What is the <u>most</u> important factor in determining your final grades?

a. Your own effort

b. Your GCSE score

c. Facts and information given by the teacher

d. Amount of marking done by the teacher

Please explain your answer:

2. How far do you agree that you can confidently identify strengths and weaknesses within your own work?

a. Strongly agree
b. Agree
c. Disagree
d. Strongly disagree

Please explain your answer:

3. Please circle which option you think conveys a teacher's <u>most</u> important function:

a. To tell you how well you are doing

b. To provide you with facts and information

c. To help you become an independent learner

Please explain your answer:

4. **Please prioritize which responses best fits your feelings about writing an essay for homework where 1 is the best fit and 5 is the least:**

Response:	Number 1-5, where **1** is the response which **best** describes your feelings
You have an opportunity to show the teacher what you know;	
You can explore an issue in more detail;	
You can practice and make mistakes;	
You feel nervous and/or worried;	
You can find out what grade you are currently working at	

Please explain your answer:

5. **Please prioritize which response best fits your reaction to receiving teacher feedback on your written work where 1 is the best fit and 5 is the least:**

Response:	Number 1-4, where **1** is the response which **best** describes your feelings
You look at the grade and use that to decide whether you are currently doing well or not in History	
You read through the comment and look to see how it applies to your work	
You read through your comment and your work and consider how to apply it to your next piece of work	
You look at the grade and comment and feel an emotional response (such as very pleased or upset)	

Please explain your answer:

6. **Do you think your reaction differs when you have self-marked and given yourself some feedback, and if so, in what ways?**

Please tick here if you consent for the data you have provided in this questionnaire to be used in my final essay. All student and staff names and the college name will be anonymized. ☐

Example of coded data from questionnaires

Please tick <u>one</u> of the following options for each assessment method:

	Please explain your decision:
…'**diagnostic' self-marking** (highlighting key criteria in your work such as 'evidence' and 'analysis')? *Coding key:* *Accuracy (6/30)* *'see'/'identify' (20/30)* *Understanding (2/30)* *Other (2/30) (1 'second opinion', 1 prefer model answers)*	• I'm never sure if I'm highlighting correctly • It's useful as it allows you to see your own mistakes first hand and also makes you able to accurately pick out specific problems ('see') • It is useful to see what you did well but where you can improve • Able to see more clearly the key aspects your including consistently and others you're falling short on • Helps define terms like 'analysis' etc. • Allows me to see what I have done well • Bias marks awarded in favour of yourself • Highlighting makes it clear and memorable • Helps to see where I did well and how to improve • Helps me to see where I need to improve • It helps to identify where I need to improve • Sometimes my mark isn't very accurate • I can point out what im doing and recognize when making a mistake (sic) • I like looking through my work in depth and analyzing the language I've used • Helps to identify analysis, evaluation, etc. • Good but want it checked over by teachers • Helps us identify our good points and ensure we repeat the 'goods' in future works • I find it more useful looking at excellent examples so I can see where I'm going wrong. • Biased- may give yourself a higher achievement level than the work actually is • Helps me to see where I have included certain things. This can also help see where I'm missing things • Clearly laid out • See what to improve • Useful to see criteria separately and see what is missing • Allows me to personally see mistakes • I usually know which is part of each criteria, but it does help me see how much of an overall essay is including key criteria • As it means I am able to read through my work and therefore spot some mistakes and correct them • A lot of the time I don't know how I am meant to mark • I have a fairly negative view of my work so going through it tends to prove fruitless as I don't see much evidence of key criteria. Also, sometimes I misunderstand the criteria so going through my work can give false confidence. • Useful because I can see what needs to be added • Useful reflection to pick out good points in own work and to highlight areas for improvement

List of indicative questions for group interviews

You will be asked a series of questions about the History department's formative assessment policy this year. The following is a list of indicative questions that might be asked:

- Did you find it helpful going through your work before you handed it in?
- Do you think you pay more attention to your own comments or the teacher's?
- Have you found using model answers helpful?
- Do you go back and check your previous feedback before you start the next essay?
- How important is teacher feedback?
- Is there anything we can do to encourage people to make better use of feedback?
- Why do you think you are encouraged to self-mark your work?

Extract from transcribed interview

Interview 3

JW: So I understand that there has been an effort to begin to get you to self-assess your work, to look at your work before handing it in, to make comments on it, etc.

General agreement.

JW: Do you find going through your work in class before handing it in a useful process?

ED: Personally I hate it.

RS: Yeah, same

ED: I can't really think to explain why

TE: I guess you're just kind of a lot more critical of yourself…

EC: Yep.

TE: …so your kinda like you point out flaws that technically aren't really there or don't really matter in the large scheme of things.

EC: Or you just can't tell whether you've done something right or done something wrong.

RS: Yeah, like you need the expert advice

JW: Sure, sure, um, and what about yourself?

ED: Yeah, I think it is easier when we are given, I found it quite useful when we were given like sample students' work. I find that quite useful. I don't know why that's different but I found it more useful to mark that than to…

JW: And are those always model answers, ideal answers, or are they sort of a mixture?

ED: It's a mixture actually which is really helpful because it's not always helpful just to see the, sort of the ideal work. You got to see both.

JW: Yeah, do you think that in these processes you are beginning to learn more of what a good essay looks like and what makes a stronger essay?

EC: Is this just for History?

JW: For History, particularly, or indeed other subjects in fact, I am quite happy to hear about other subjects and different approaches

Extract from transcribed interviews with analytical commentary

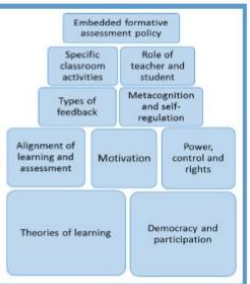
Interview 3 transcript:	Analysis:
JW: Sure, it's not a standard formula. ED: Exactly. TE: But, but that's kind of my problem with um, with model answers like if it doesn't, **I feel if it doesn't align with that of the model answer I've done something wrong** (EC agrees) or **I feel really depressed because I haven't done it to this standard or I haven't structured it in this way** EC: - and you feel **too much pressure to kind of conform** to that mark scheme. TE: Yeah, and well while I still might get the same marks **I don't personally feel as if it's anywhere near as good** as this one **so therefore I've completely failed** so… *JW: It sounds like you're quite harsh on yourself as a marker* TE: Yeaaah (small laugh), **I'm pretty harsh**, so, you know (small laugh). *JW: And is that true of you as well?* ED: Yeah	This differs to RS's earlier comment that looking at model answers are helpful. Instead, these comments seem to show model answers as a source of stress – *'I feel really depressed'.* '*in this way'*- Implicit to this comment is the idea that there is one set way to write an essay? If there was a strongly held understanding that there are a range of ways to write a quality essay, this might be less of a concern. EC makes this more explicit – *pressure to…conform'* – these are both higher achieving students. Do they feel limited by model answers? This is an interesting comment. It suggests that the grade is not the indicator of success – *I still might get the same marks* – if the student were only concerned about the grade, then the differences in essay style would not perhaps concern them. Instead, they are clearly very bothered by the perceived difference in standard – *'completely failed'.* The fact that this student has a judgement of quality beyond the grade does suggest some degree of mastery learning – they want their essay to be the best – but on the other hand it is clearly a source of stress for them. *'I'm pretty harsh'* – again, highlights the emotional response that can be involved in self-marking.
JW: Yeah so that's part of, what makes you uncomfortable is it that you're being, that you're looking and thinking, 'oh this is rubbish! I'll throw it in the bin'. What about yourself? GS: Well I don't know like because sometimes, because obviously **we want to be self-independent for the future**, and if we are **looking at model answers and we have to try and like, like, copy those in a sense**, err so it kind of backs off, it **doesn't make like for self-independence, you can't be free**, you can't be like free to write… *JW: …If you are imitating somebody else, yeah.* TE and EC laugh and agree GS: It's a **bit limited** in what you can actually do. *JW: Yeah, so did you learn to, do you feel you've learned to make comments on your own work or to see the strengths and weaknesses in your work?* GS: Err**, I have been trying** but I think once again um, there might be flaws in my work which I, **I might think is amazing but then it might not be so good after all** (laughs).	GS highlights a concern that model answers are limiting because they restrict independence. He makes an interesting point – by highlighting the fact that this method does not help students become independent, it would suggest that this view is underpinned by the assumption that formative assessment ought to make students independent – or at least that is an aim of education more widely. Therefore, although this student recognizes the problem with achieving this through this particular process – it is interesting that he criticizes it for this reason, thus revealing the importance placed on becoming independent. *'bit limited'* – again, the limitations of model answers. *'I have been trying'* might indicate the difficulty the student finds self-assessment (or this could be an attempt to say that they have simply not found it effective, but wanted to 'soften the blow' by saying they had tried). *'it might not be so good after all'* – suggest concerns of accuracy of self-assessment. This male student's comment differs from female student in interview 1 who was more concerned with using self-assessment modestly (gender difference?)

Research poster (used at MLT conference and with college TLC)

Using student views to develop a formative assessment policy in a sixth-form History department

Literature review:

The literature review was the starting point for developing my project. The literature on student voice and democratic classrooms encouraged me to design a methodology where **student voice** was central so that the outcomes of the research would not only be of use to developing my teaching practice, but the process of the research would also empower the students. I divided my approach to the literature on formative assessment into 'big pictures' and 'little pictures'. The 'little pictures' focused on the practical use of **formative assessment in the classroom**, what Black and Wiliam describe as a *'kit of parts'* (1998:38) and I used this to develop my assessment policy for this year. My 'big picture' looked at the wider purpose and aims of formative assessment, including the role it has in **empowering students** and shifting the power balance of the classroom. I developed my research questions out of the literature review and my planning of the methodology with the aim of assessing how effective my use of formative assessment had been this year, and how to improve it next year.



Research questions:

1. Do the students report more engagement with self-assessment compared to the students I taught last year? (Questionnaire)
2. To what extent do the students consider current self-assessment strategies to be effective? (Questionnaire)
3. Do the students report greater feelings of responsibility in the assessment process? (Questionnaire and group interviews)

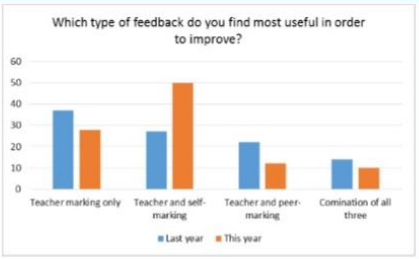
Methodology:

The intervention I applied this year was to make **self-assessment an embedded part of the assessment process**. All essays would now be self-marked before submission and students would have an opportunity to draft a paragraph one week before an essay is due to get some self and peer feedback. I then aimed to collect data from the students to assess how effectively this was working, and what could be changed for next year. This led to the following steps:

1. Questionnaire—using the same questions from my part 2 questionnaire to see if students reported more confidence in self-assessment now that it was applied regularly.
2. Questionnaire—to gather student views on which specific formative assessment activities they felt were most useful
3. Questionnaire and group interviews— to explore student confidence and independence in the assessment process. This would help me judge the effectiveness of my intervention, and help me understand which attitudes still needed support.
4. To redraft the department's assessment policy, and give to classes to review in a classroom activity.

Initial findings:

I have found that in comparison to my students last year, this year double the number of students reported greater confidence that self-assessment (when combined with teacher assessment) helped them to improve. However, support for self-assessment without teacher assessment remained slight.



Evaluation of initial findings:

Although these findings suggest that my intervention has been successful, it does raise further questions. Do my students report greater confidence in the ability of self-assessment to help them improve because they have become more independent learners, or am I observing a 'learned behaviour' whereby a practice that I have told them is beneficial is now being displayed? I hope that my third questionnaire into student feelings of responsibility and agency will help me explore the extent to which this is a 'learned response' or evidence of a genuine increase in student agency.

Problems encountered:

Designing my research project led to many dead ends. Approaches that had to be abandoned or changed included staff interviews and the timing of my data collection. However the most significant problem encountered was in narrowing down my research question. The main 'dead end' here was trying to prove that my formative assessment interventions had worked. As Wiliam highlights, it *"cannot be proved that formative assessment did improve learning as that would require the counterclaim that what happened would have been different to what would have otherwise happened"* (2011: 43). To work through this problem, my main focus became using the data to improve and redraft our formative assessment policy. This still relied on some evaluation of what was working, but now focused on students reporting their confidence in each measure and using their views to improve the policy for next year and therefore giving my research a more practical outcome.

Key references:

Dann, Ruth (2002) *Promoting Assessment as Learning, Improving the Learning Process.* Routledge Falmer, London
Sadler, D.R. (1989) Formative assessment and the design of instructional systems, *Instructional Science,* 18, pp. 119-144
Torrance, Harry and John Pryor (1998) *Investigating Formative Assessment.* Open University Press, Buckingham
Wiliam, Dylan (2011) *Embedded formative assessment*. Solution Tree Press, Bloomington

Emma Hawkins *emma.hawkins@kellogg.ox.ac.uk*