

LEVEL 7 ASSESSMENT SPECIFICATION

Student name:	Leah Jane Williams	Student ID number:	1806066
Programme:	Secondary PGCE Geography		
Module:	Curriculum Development and Enhancement		
Module code:	ECSP7002	Contribution to Overall Module Assessment (%):	100%
Lecturer:	Helen Denny	Internal Verifier:	
Assignment Title:	Research Study including Ethics Form	Word count (or equivalent):	10.000 words in total including an Ethics form (to be given in on January 11 th 2019)
Submission deadline:	May 13 th 2019	Return date of provisional marks & written feedback:	June 10 th 2019
Submission method:	All written assessments, where practical and possible, must be submitted via Turnitin unless otherwise instructed by the Lecturer. (Please DO NOT put this assessment specification into Turnitin or it will match many similarities with other students' submissions.) Alternative submission method (if applicable): Late submission of the assessment will result in a late penalty mark. Penalties for late submission: Up to one week late, maximum mark of 50%. Over one week late, 0%. Only the Extenuating Circumstances Panel may approve a change to submission dates.		
Academic honesty / referencing:	Academic honesty is required. In the main body of your submission you must give credit to authors on whose research and ideas your work is based. Append to your submission a reference list that indicates the books, articles, etc. that you have used, cited or quoted in order to complete this assessment.		

This research was carried out with Class A from School X. School X is a mixed gender comprehensive comprising of 11 to 18-year olds within the local authority of Bridgend. As of 2018, 1214 pupils are on roll with an average attendance of 92.8%. The school sits above the national average of 17.4% Free School Meals (FSM) (Estyn 2015) with a total of 24.8% of pupils receiving FSM at School X. The school is in the Amber support category. Class A is a mixed gender class of 12 females and 8 males (20 pupils) aged 11 to 12. In order to maintain anonymity participants were assigned a number and referred to as Pupil 1, Pupil 2 etc.

From the time spent at School X and with Class A, it became apparent that although commented on in both the departmental improvement plan (DIP) and school improvement plan (SIP), rarely was true self-assessment carried out as the focus was on peer and teacher-based assessment. Similar observations can be made on the use of success criteria which seemed to be reserved for specific end of topic assessments, rather than being provided and used for extended pieces of writing in the classroom or home. It is for this reason that the author decided to carry out research into the use of self-assessment and success criteria as an effective tool for assessment for learning (formative assessment).

This research is focused on the mechanisms of formative assessment (assessment for learning), derived from Blooms (1968) model of mastery learning, distinguishing between formative and summative assessment in the classroom (Bloom, Hastings and Madaus, 1971). The majority of the literature focused on assessment for learning (formative assessment) is based on the work of Paul Black and Dylan William. They defined formative assessment as *“all those activities undertaken by teachers, and/or their students to modify teaching and learning activities in which they [the students] are engaged”* (Black and William, 1998; 8). From this definition confusion can be seen in the literature with the term “formative assessment” being misunderstood and confused by some (James et al., 2006), resulting in limitations on its effectiveness in learning outcomes (Coe, 2013; 10). The Department of Education (DoE) reports that *“formative assessment was not always being used as an integral part of effective teaching”* (DoE, 2015; 13) and teachers were simply tracking pupils progress towards target levels due to a focus on National Curriculum levels rather than focusing on identifying *“strengths and gaps in pupils’ knowledge and understanding of the programmes of study”* (DoE, 2015; 13). Hattie and Timperley (2007) sum up these approaches as methods of identifying gaps in pupils’ current knowledge and where they need to be, providing *“steps that will reduce or close the gap”* (Hattie and Timperley, 2007; 102). The available literature raises the important point that links between formative and summative (assessment of learning) assessment must be drawn by teachers and pupils. Some students prefer the use of summative indicators to reflect on their own learning (Brookhart, 2001) and teachers can make formative use of summative data (Black et al. 2003) to enhance reflection.

This emphasises the need for the self-reflection of pupils to help them thrive and succeed, rather than simply act as a data set. Two of the main influences that can impact assessment for learning are self-assessment and the use of success criteria.

Monitoring and evaluating one's own performance is key to learning, not only in a classroom environment but throughout life, therefore it is vital that pupils are equipped with the ability to carry out self-assessment. Self-assessment is a key element in assessment for learning as it involves pupils in evaluating the quality of their own individual work rather than a reliance on teachers for a sole judgement. Research suggests that using pupils to evaluate their own work is a useful tool in the development of education (Andrade & Boulay, 2003; Andrade, Du, & Wang, 2008; Ross, Rolheiser, & Hogaboam-Gray, 1999). Donaldson (2015) comments on its importance in helping pupils take responsibility for their learning *"by reflecting on their understanding and performance in relation to clear criteria, they can identify areas of difficulty that might not be easily seen by an external observer"* (Donaldson, 2015; 79). This notion is reflected by Black and William (2006) who deem self-assessment, alongside peer assessment, as *"essential to learning because students can only achieve a learning goal if they understand that goal and can assess what they need to do to reach it"* (Black and William, 2006; 15) and argue it is a useful strategy in raising academic performance (Black and William, 2006). It is apparent to the author that self-assessment should serve as a solid foundation in assessment for learning and teachers should share their learning intentions with pupils (Clarke, 1998). Its validity and effectiveness as a method of assessment however varies (Brown and Harris, 2013). Goodrich (1996) believes that necessary conditions must be in place for pupils to carry out effective self-assessment:

- Awareness of the values of self-assessment
- Access to clear criteria on which to base the assessment
- A specific task or performance to assess
- Models of self-assessment
- Direct instruction in and assistance with self-assessment
- Practice
- Cues regarding when it is appropriate to self-assess
- Opportunities to revise and improve the task or performance

(Goodrich 1996; 16-17; Wiggins, 1990)

It must be acknowledged that not all of the above criteria may be met, but an importance on awareness, clear criteria, modelling and the appropriateness of self-assessment is key. Cardinet (1977) proposed that formative self-assessment acts as a regulator for student learning through integrated guidance in each activity, further supporting Goodrich (1996) suggestion of modelling and specific tasks.

Although the benefits and importance of self-assessment are clearly noted in the literature, it must be an effective practice. Self-assessment is only beneficial when it serves a purpose in the curriculum and not simply a 'tick box exercise'. It must also be noted that pupils cannot be expected to carry out self-assessment without being given the correct tools and guidance from teachers. The focus of this form of formative assessment is *"helping the teacher identify areas where more explanation or practice is needed"* (Assessment Reform group, 1999; 7), but more importantly in

aiding the pupils to identify areas of improvement themselves without focusing on the marks/levels given; a fixation often associated with the term 'assessment'. William highlighted this issue during a 2012 interview published in the Times Educational Supplement, where he commented on "*the big mistake*" (Stewart, 2012) of using the term assessment due to its association with exams. In later literature William (2013) goes on to suggest the term "*responsive teaching*" (William, 2013) as an alternative term. Following the conditions put forth by Goodrich (1996), for self-assessment to be deemed an effective practice it must be understood, facilitated appropriately, repeated and acted upon. It is the later, the need to act upon identified targets, which is key. Without the opportunity for pupils to carry forward and revise their targets then they cannot improve and work towards where they need to be.

Pupils are constantly evaluating and assessing themselves and their peers, whether they are aware of it or not. By giving the process a name and a route to follow (via modelling self-assessment) the pupils will be able to be more focused on their work, with the hope of proving beneficial with higher levels of attainment being reached.

A detrimental effect on pupils may be caused by the use of self-assessment. Pupils could come to wrong conclusions regarding their ability and either overestimate or underestimate what they are capable of. Human nature can lead to unrealistic optimism about ability, with potential for the overlooking of key information resulting in deficits in knowledge and skills (Dunning, Heath and Suls, 2004). On the other hand, pupils who underestimate their ability could begin task avoidance and disengagement with the subject, damaging not only their education but potentially their wellbeing (Ramdass and Zimmerman, 2008). It is for reasons such as these that teachers must facilitate the use of self-assessment in the classroom and have an open dialogue with pupils in order to identify progression steps. It also highlights the need for praise within a classroom to maintain the want for learning and progression in learning. These issues also raise questions regarding self-assessment as a stand-alone tool, suggesting that it must be used in conjunction with traditional teacher led marking and peer assessment.

The use of success criteria is closely linked with the use and effective use of self-assessment. How is it possible to expect pupils to reflect upon their work if they do not have the structure or model from which to evaluate? Modelling success and demonstrating what it looks like is a key aspect of assessment of learning, one pupil has described it as "*like knowing the teachers secret*" (Spendlove, 2009; 18). The regular use of success criteria also opens the door for a greater level of teacher-pupil dialogue, allowing for oral feedback regarding a pupils next steps in reaching where they need to be, specifically how they get there. The process need not be formal in nature, simple conversations where teachers listen to their pupils can be just as effective, as long as the pupils know where they currently lay and the actions, they need to carry out to move forward and progress.

Although the benefits of these forms of assessment for learning are well documented, they cannot be used as standalone methods. Sadler (1989) notes the importance of any kind of feedback from informing teachers about understanding,

facilitating learners in the identification of gaps, the selection of suitable tasks and the modification of the teacher and teaching in order to support pupils (Sadler, 1989). It is the facilitation of teachers to support pupil progress which is key. Time must be given in order to understand how the process of self-assessment is carried out as well as modelling the use of success criteria. This may be carried out (as seen in this research) by developing a success criterion as a class, working through a piece of work together, or embedding the use of self-assessment into the curriculum so it can become second nature to pupils and routine practice throughout their education.

It is apparent that it may not always be possible. Pressures may exist on teachers and pupils to produce high levels of attainment in the form of levels and results, resulting in (potentially) the conscious neglecting of formative assessment in favour of a series of summative assessments to meet data requirements (Fautley and Savage, 2008). There is still vast room for the improvement and development of assessment for learning, in particular self-assessment. Bennett (2011; 23) makes the point that it is still a “work in progress” and that expectations and claims surrounding the topic must be sensible and scrutinised in order to “realise its considerable promise” (Bennett, 2011; 23).

Self-assessment (as well as other formative methods) can be integral to effective teaching on a day to day basis (Leahy et al., 2015). Although noted in both the school and departmental policy (alongside both the school and departmental improvement plans) little to no use of self-assessment in practice was observed during the authors time at School X, with peer assessment being the preferred method of pupil lead assessment for learning. The policy states that at least one task is to be self-assessed by pupils each term using purple pen (to distinguish from teacher assessment/marking carried out with green pen) and one task to be peer-assessed by pupils each term. The only self-marking observed by the author was its use to mark small question tasks, with pupils ticking or crossing the correct answer from a series of questions given, usually as a starter activity. No self-assessment with the purposes as discussed above was carried out, therefore providing no real purpose or opportunity of progress for learners. As mentioned, the focus was on peer-assessment which is carried out for every end of topic assessment, known as “write on” and “think on” tasks. Within the Geography department of School X there is a focus on “write on” extended writing tasks, often in the form of an argument based written task (e.g. newspaper article or formal letter) where pupils give their response to a given statement. It is school policy that all pupils are given the opportunity to complete “reading habits” relating to the assessment (a strategy devised to allow pupils to read and analyse text ensuring understanding whilst helping to improve their comprehension and critical thinking skills). A rubric is also provided for every write on task as a success criterion for pupils to follow and use to model their answers. Rubrics are a constant throughout all year groups, the idea being that once explained in year 7 pupils are equipped with the ability to use them throughout their education. As with self-assessment, the use of rubrics (and any success criteria or answer modelling) is limited to these end of topic assessments, which according to the literature should be common practice for all marked work (when

appropriate). What is key with the policy of School X is the use of “feed forward” lessons. Once the assessment has been completed, it is first peer assessed using the same mark scheme as the class teacher, and then pupils are given the opportunity to re-draft elements of their work based on their peers comments (in the form of what went well (WWW) and even better if (EBI)) before the final assessment level is given by the teacher. This use of peer assessment places an emphasis on praising pupils for their improvements. This can be emulated with the use of self-assessment, as long as it follows the principles detailed above, providing pupils with an effective tool to further their own ability and education.

Before carrying out the intended research, all members of Class A were briefed on what the research entailed, how themselves and related data would be anonymised and protected, whilst given the option to withdraw participation at any time without consequence. The informed consent of individual participants was key. All pupils wished to participate, providing the author with a data set of 20 pupils. Throughout the duration of the research observations were carried out in order to observe participants before, during and after the completion of the stages discussed below. The validity of the research was influenced by the knowledge gained through observation as issues encountered by participants were identified, and comparisons drawn between the pre and post diagnostic stages.

Following some suggestions put forward by Leahy et al. (2015), the aim of this research was to investigate the impact (if any) self-assessment and target setting can have on an individual’s extended written work, notably improvement in level attainment and progression overtime. In order to investigate this overarching theme, 3 research aims were developed:

- 1) To explore the impact success criteria has on written work
- 2) To investigate the use of participants setting personal targets for their own improvement
- 3) To make recommendations based on the impact of self-assessment can have on participants and its usefulness

The research itself was divided into stages: pre-diagnostic, intervention and post-diagnostic. The pre and post-diagnostic are similar in nature both being extended pieces of writing, however due to the constraints of the scheme of work they fulfil different purposes. The pre diagnostic test was a diary entry with an emphasis on empathetic writing, based on work carried out during the term. The post diagnostic test saw pupils write a letter, employing persuasive techniques and information learnt during the term. Both tests were ‘open book’ in order to remove the stress of having to recall facts at short notice, especially as the idea behind this research is to identify targets for improvement and self-assessment, not to test subject knowledge. Saying this the tasks proved useful in consolidating learning at the end of the topic.

By beginning the research process with a pre-diagnostic test, the author was able to identify the current standard of written work demonstrated by Class A (without the use of a success criterion). Alongside establishing the base level for research, the

author was able to observe how well Class A responded to a success criterion to evaluate/critique their work, as well as their ability to self-assess during the intervention stage after the modelling of self-assessment by the author. This process gave participants scope to identify their own targets with regard to their written work.

Once the initial pre-diagnostic writing task (a diary entry based on what was learnt in the previous lesson) was complete the research moved into stage 2, the intervention. This stage was focused on the participants refining their self-assessment skills, acting on their individual targets identified through self-assessment and producing a success criterion to carry forward into the post-diagnostic test. The intervention stage was carried out over a two-hour period (two lessons). During this time pupils reflected on the extended writing task completed during the pre-diagnostic stage, identifying areas of improvement, setting targets and developing a success criterion (based on targets) as a class to be carried forward into the post-diagnostic stage.

Once the intervention stage was complete Class A moved into the post-diagnostic stage. During this stage participants used the success criteria generated during the intervention stage to carry out another extended piece of writing similar to that used in the pre-diagnostic stage, in this instance a letter. The key emphasis was to utilise the success criteria and the identified targets of the participants.

Due to time constraints and a limited number of classes being available to participate, the Class A sample was recruited through convenience sampling. Although simple and can easily be carried out in a short period of time, convenience sampling can lead to high sampling errors (data not representing the entire population, in this case of School X) and potential selection bias (lack of randomisation of participants). The short period of time available to complete this research may also hinder the process as the intervention may be considered rushed and the full potential of self-assessment and target setting not fully explored by both the teacher and participants.

At each stage the author assessed participants work to see if the targets and levels determined by themselves matched those the author would assign. By doing this any misconceptions and misunderstandings could also be addressed and common targets for improvement identified. Assessment levels were graded at level 3, 4 and 5 aligning with school policy and National Curriculum levels for Key Stage 3 (KS3). Each level was subdivided into A, B and C which would determine where in the level the participant would lay, A being a 'high' level and C 'low' level. Two levels would be produced for each test with an average being taken for the overall level for that piece of work. One level will focus on subject knowledge and the second will focus on literacy (as per school and departmental policy).

The design of this research can easily be replicated to encompass a larger number of participants in order to gain a more accurate representation of the School X population.

The four hours/lessons assigned to carry out this research with Class A proved effective, but evaluation and reflection is greatly needed.

15 of the 20 participants from Class A produced a different level in the post-diagnostic test following the intervention stage to those produced in the pre-diagnostic test. In the majority of cases (14 pupils) progress was made with either increases in sub or full levels. One pupil showed a decrease of one sublevel. Gender was considered as a potential influence on the results of this research, however on inspection no clear pattern emerged with regard to attainment and progress so it was not included in the results. However, increased sample size may allow for more focused analysis on the relationship between gender and attainment to be carried out in the future.

The pre-diagnostic test produced a range of levels throughout Class A from 3c to 5c. 7 (35%) participants were at level 3, 11 (55%) participants level 4 and 2 (10%) of participants at level 5. This reflects the mixed ability of the class. The post-diagnostic test again produced a range of levels throughout Class A from 3b to 5c. 4 (20%) participants were at level 3, 12 (60%) participants level 4 and 4 (20%) of participants at level 5. When broken down further it can be seen that 1 (5%) participant improved by a full level, 7 (35%) participants by 1 sublevel, 6 (30%) participants by 2 sublevels, 5 (25%) participants demonstrated no change in attainment and 1 (5%) participant dropped 1 sublevel (Table 1, Appendix A). From this data it can be said that through intervention and the appropriate scaffolding and modelling, self-assessed target setting, and the implementation of success criteria can increase attainment and pupil progress. However, these results only reflect a very small proportion of the School X cohort and cannot be treated as concrete evidence.

This research was carried out over the space of four lessons, a total of four hours class time. One hour was used for the pre-diagnostic test, two hours were allocated to the intervention stage and one hour to the post-diagnostic test. This sequence was designed to allow for the clear progression of participants, with each stage building on the previous. It should also be noted that as no specific readings were required to complete the tasks involved in this research, a specific reading habit (as detailed in School X's policy) was not completed nor necessary.

The pre diagnostic test was used by the author as a base level test to gauge pupils extended writing skills following basic instructions (to write a diary entry based on knowledge gained in previous lessons, focusing on the use of empathy). Apart from instructions detailed in the task brief, no specific success criteria or rubric was available to pupils for this. Participants were given the full one-hour lesson to complete the task. The use of the full hour allowed pupils to read and reflect on their work before completing the self-assessment during the intervention, showing that pupils are constantly evaluating without prompt. It also allowed pupils to focus on the written literacy elements of the work.

Once the pre-diagnostic stage had been completed the intervention stage could commence. The first lesson dedicated to the intervention stage was focused on explaining the process of self-assessment, self-assessing the pre-diagnostic test and

answering any questions regarding the tasks. Participants were given a basic success criterion and rubric (Appendix B) and told that they would be assessing their own performance against these. Before beginning the process of self-assessment, it was made explicit that participants must read through the criterion in order to highlight any issues and misconceptions. As a way of modelling self-assessment for participants the author worked through an example piece of writing similar to the pre-diagnostic test but unrelated to the work and participants of Class A. By visualising self-assessment as a class, participants were able to understand the process before carrying it out themselves. Once all questions had been answered and participants understood what it was, they needed to do, the self-assessment took place. It was made explicit to participants that the main goal of this was to identify targets for improvement, not the overall level generated through “assessment”. Once the self-assessment had taken place and pupils had identified their own targets, their work was swapped with other members of the class for peer-assessment. By carrying out peer assessment in conjunction with the initial self-assessment allowed pupils to see if the views of peers aligned with their own, opening up the potential for further target identification.

The second lesson of the intervention stage focused on developing a success criterion to be used in the post-diagnostic test, another extended piece of writing utilising knowledge already gained to write a formal letter using persuasive language. The starting point of this lesson was the exposition of the post-diagnostic task, the writing of a formal letter using persuasive language. By explaining the task first, participants would be more focused in the production of success criterion for the task. Once the task had been explained and questions answered, Class A returned to their now self (and peer) assessed pre-diagnostic test. As a class all targets were pooled. To do this the author systematically went through the self-assessed work to view the targets participants had identified (as well as marking the work their self), collating them and anonymising them so Class A would not know whose target was whose. This also allowed the author to identify common targets for improvements which would allow for a more direct and focused period of intervention. Once the initial targets had been identified they were presented to the class. Using the task brief given at the start of lesson and their targets pupils were tasked to create a success criterion to use while completing (and subsequently marking) the post-diagnostic test (Appendix B). It should be noted that due to school and departmental policy targets for literacy must be included as standard. This meant that participants decided to have a separate section of the success criterion dedicated to literacy requirements (Appendix B).

By completing the intervention in this order, participants were first walked through the process before carrying it out independently. It also ensured that the appropriate modelling and scaffolding had been provided by the author, so participants knew what was required and how to proceed. What was key to this stage was making it know to pupils that the important part of the process of self-assessment is the identifying of their own targets in order to progress in their learning, not the levels in which they (and the author) assigned them.

The research process concluded with the completion of a formal letter employing subject knowledge from previous lessons and using persuasive language. This post-diagnostic test was designed to allow pupils to carry forward their targets and their new knowledge regarding the use of success criteria and self-assessment. The same conditions were applied to this as the pre-diagnostic test, with the exception of participants having a success criterion available to them. By using a success criterion when completing the task, pupils knew exactly what needed to be done in order to achieve specific levels, hopefully beyond their target level, mirroring the comments made by pupils in Spendlove (2009; 8).

Due to a number of factors beyond the authors control, a limited number of participants were available to take part in this research, heavily influencing its design and method of this research. There were also constraints from the department scheme of work as certain tasks needed to be carried out during the research window. These tasks were both extended writing tasks focusing on emotive language, the use of persuasion and empathy, therefore creating the pre and post diagnostic test tasks and conditions. Although similar these tasks have a different motive and do not fulfil the same purpose. If repeated the pre and post diagnostic test need to be much similar in nature to gain a better representation of what can be achieved through self-assessment. Rather than implementing a success criterion, which although implemented in this research can be time consuming, a simple knowledge test can be used in both instances with a different assessment for learning strategy employed, such as the traffic light system for an instant assessment of how pupils feel about a topic. There may also be a bias involved as the author is the class teacher. In order to minimise bias assessment should be carried out by a teacher outside of the classroom and research.

In conclusion, this research shows that target setting through self-assessment when used with the appropriate scaffolding can indeed improve the extended writing skills and overall level for the majority (70%) of participants in Class A, a pattern demonstrated in current literature regarding pupil self-assessment and target setting.

However, these results are a reflection of only one set of individuals and cannot be conclusive proof that the implementation of success criteria and self-assessed target setting works in all cases.

What the research detailed above has shown, although its conclusions are not concrete, is that extensive training in self-assessment is not required to see results in pupils progress, with 14 out of 20 participants showing an increase in attainment after one to two hours of modelling and 'training'. With this being said sufficient time needs to be given to maximise the benefits. In most cases no significant increases in attainment have been achieved by the participants of Class A, with few increasing their grade by full levels (e.g. level 3-4) and the majority increasing by sub levels.

This type of formative assessment must be carried out with great thought and planning in order for its potential to be maximised. Pupils cannot be expected to carry out self-assessment without an open dialogue with teachers, facilitating

understanding and aiding pupils in taking responsibility for their own learning. Although important, the onus cannot be fully placed on the pupils and teachers must be held responsible for helping pupils in recognising and reaching the progress they need.

This research was only focused on the 20 pupils of Class A. Although conclusions can be made (mainly that the use of success criteria can improve the level of pupils' work), they are limited. 20 pupils are not a suitable sample size and this study must be expanded to other pupils in order to provide a truer representation of its success, or failure. Other classes of a similar demographic should be studied, as well as different demography e.g. age group in order to see if the findings are apparent throughout the school or just individual class patterns. By increasing the sample size and demographic, further conclusions may be drawn regarding the age and gender of pupils throughout School X, as well as comparisons between sample groups. Increasing the study size will also allow for the identification of individuals who would benefit most from the use of self-assessment and target setting, highlighting any significant changes in attainment. As mentioned above, increased sampling size may also allow for investigation into the relationship between gender and attainment.

The time allotted to the intervention stage may need to be reconsidered in future research. Although the two-lesson period provided some results, a longer period of time should be utilised in order for greater exploration of self-assessment and target setting methods, as well as the development of practises to be embedded in day to day classroom assessment.

Serious time constraints were placed upon this study and due to numerous activities, that pupils within Class A participate in meant that not all pupils were available for the full four hours. All completed the required tasks and were instructed on how to complete self-assessment but added time pressures may have had an impact on the research. The timed nature of the pre and post diagnostic tests could also have had an impact on the performance of the participants. This could be a variable introduced in future studies. By following a similar structure and using pre identified targets, tasks could be completed under different timed conditions in order to see if this affects attainment.

Pupil voice should be taken into more consideration, giving participants more freedom in discussing and developing how they wish to assess their individual work, the format in which they wish to carry out the assessment e.g. an assessment sheet, in combination with peer assessment or [as carried out in this research] developing and utilising success criteria. Pupils understand themselves and how they work better than any teacher or researcher. It is important that they are given the freedom and tools to discover how best to improve their educational experience and thrive in the class room. Pupil voice may also be used to self-assess and reflect in every lesson. By incorporating strategies into everyday classroom activities to check for understanding and gaps in knowledge directly from pupils themselves. Some strategies observed by the author include exit tickets where pupils must answer a series of simple questions in order to ascertain what has been learnt confidently,

what needs improvement and what has not been understood; the traffic light or RAG (red, amber, green) strategy works on the same principle as exit tickets with pupils grading what they have learnt in the lesson as red, amber or green based on their confidence (green being most confident and red least). The use of the traffic light system can be developed by asking pupils to 'prove it' and provide something learnt from the lesson, or by posing a question about something they feel they need more help with. These strategies are both simple to understand and easy to incorporate into all lessons. They may be an interesting point of research for future development.

Repeat assessment should also be carried out with Class A in order to identify further progression and if any particular benefits in attainment and progress have been achieved by participants. A 'one off' increase in sub-level/level is not sufficient enough to draw any solid conclusions. The continued use of pupil led target setting through self-assessment and success criteria needs to be implemented and monitored to increase the validity of results, but more importantly allow pupils to take responsibility for their own education and development. Following on from the initial recommendation that sample size must be increased, repeated use of self-assessment must also be carried out for all future participants to monitor progress.

The simplistic nature of this research allows great scope for development, in particular changes at the intervention stage to change the focus of research. Instead of self-assessment being carried out by participants peer-assessment can be implemented. This would allow for the comparison of the effectiveness of both formative strategies in pupil progress and attainment. Self and peer assessment may also be used in conjunction with one another, giving participants a wider scope from which to set targets in future work.

Appendices

Appendix A: Results

Table 1: Overview of results for Class A

20 pupils, 12 females and 8 males.

Pupil	Pre-Diagnostic Level	Post-Diagnostic Level	Level Difference
1	4b	5c	+2 sublevels
2	4c	4b	+ 1 sublevel
3	4a	5c	+ 1 sublevel
4	4c	4a	+ 2 sublevels
5	5c	5c	No change
6	3a	4a	+ 1 full level
7	3a	4b	+ 2 sublevels
8	4b	4a	+ 1 sublevel
9	3c	3a	+ 2 sublevels
10	4b	4b	No change
11	3b	3a	+ 1 sublevel
12	4c	3a	- 1 sublevel
13	4a	4a	No change
14	3c	3b	+ 1 sublevel
15	4b	4a	+ 1 sublevel
16	4a	4a	No change
17	5c	5c	No change
18	3a	4b	+ 2 sublevels
19	3b	4c	+ 2 sublevels
20	4c	4b	+ 1 sublevel

Table 2: Overview of results for Class A

20 pupils, 12 females and 8 males

Change in Level	Number of Pupils	Pupil Number
Decrease in 1 Sublevel	1	12
No Change	5	5, 10, 16, 17
Increase in 1 Sublevel	7	2, 3, 8, 11, 14, 15, 20
Increase in 2 Sublevels	6	1, 4, 7, 9, 18, 19
Increase in 1 Full Level	1	6

Figure 1: Graph giving an overview of results for Class A

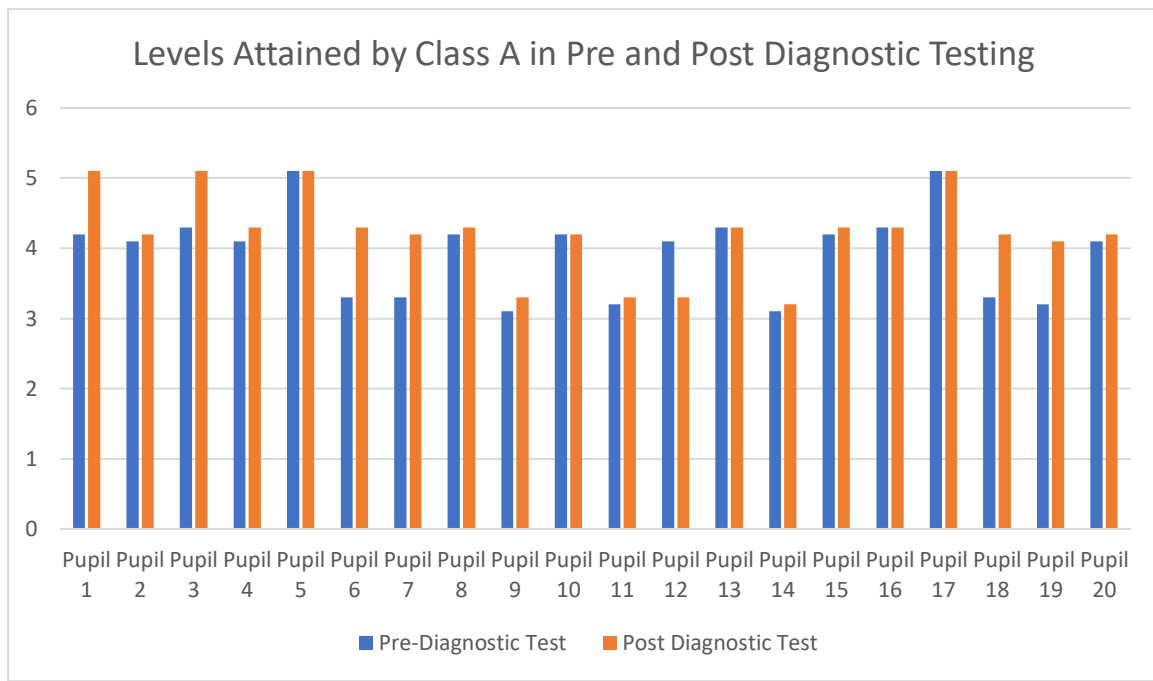
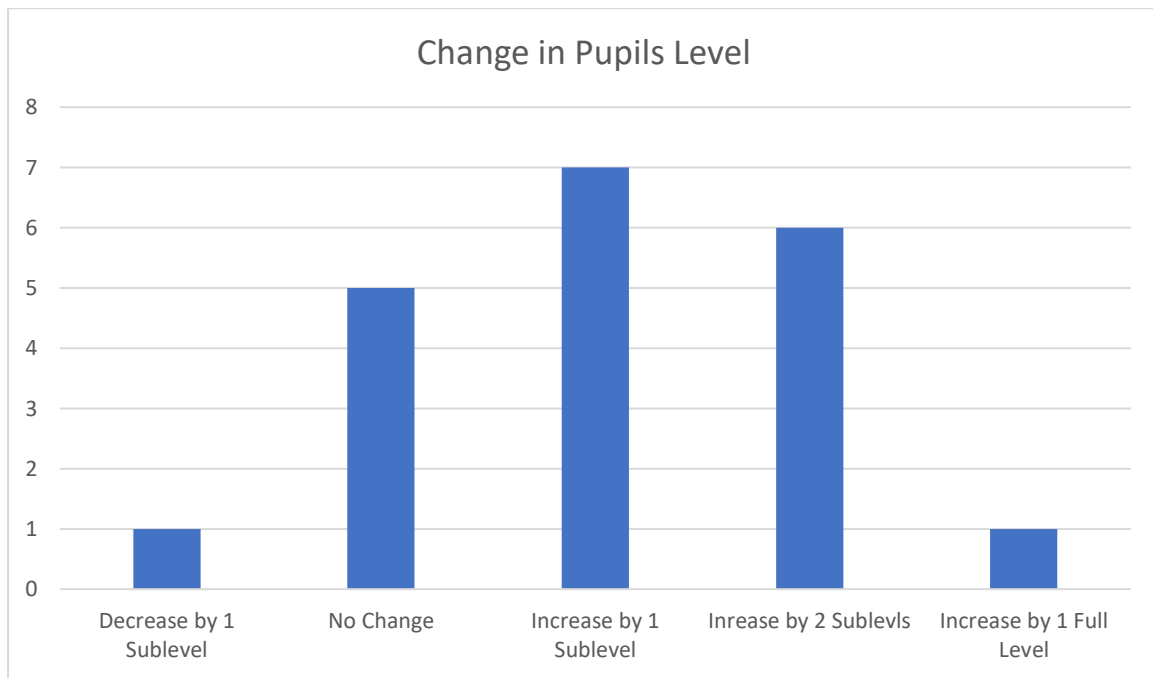


Figure 2: Graph showing the level change demonstrated by Class A participants



Appendix B

Figure 3: Pre-diagnostic task brief

Figure 4: Post-diagnostic task brief and success criteria

Bibliography

Andrade, H. and Boulay, B. 2003. Gender and the role of rubric-referenced self-assessment in learning to write. *Journal of Educational Research*, 97: 21–34.

Andrade, H., Du, Y. and Wang, X. 2008. Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practices*, 27(2): 3–13. Assessment Reform Group (1999) *Assessment for Learning: Beyond the Black Box*. Cambridge: Cambridge School of Education.

Bennett, R. E. (2011). Formative Assessment: A Critical Review. *Assessment in Education: Principles, Policy & Practice*. 18 (1), 5-25.

Black P and William D (1998) *Inside the Black Box: Raising Standards Through Classroom Assessment*. London: King's College School of Education

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

Black P and William D (2006) *Assessment for learning in the classroom*. In: Gardner J (ed) *Assessment and Learning* (2nd ed). London: SAGE, pp. 11–32

Bloom, B. S., Hasting, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York, NY: McGraw-Hill.

Brookhart, S. M. (2001). Successful students' formative and summative use of assessment information. *Assessment in Education: Principles, Policy & Practice*, 8, 153–169. Available: <https://www.tandfonline.com/doi/10.1080/09695940123775>

Cardinet, J. (1977). *Objectifs éducatifs et évaluation individualisée*. Rapport IRDP/R 77.05. Neuchatel, Switzerland: Institut Romand de Recherches et de Documentation Pédagogiques. Available: https://scholar.google.com/scholar_lookup?hl=en&publication_year=1977&author=J.+Cardinet&title=Objectifs+%C3%A9ducatifs+et+%C3%A9valuation+individualis%C3%A9e

CLARKE, S. (1998). *Assessment in the Primary Classroom*. London: Hodder and Stoughton.

Donaldson, G. (2015). *Successful Futures: Independent Review of Curriculum and Assessment Arrangements in Wales*. Available: <https://gov.wales/docs/dcells/publications/150225-successful-futures-en.pdf>. Last accessed 24th April 2019.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69-106.

Estyn (2015). *Report on School X*

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

Harvard University. 16-17. Available:

https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Goodrich%2C+H.+%281996%29.+Student+self-assessment%3A+At+the+intersection+of+metacognition+and+authentic+assessment.&btnG=

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112. Available:

<https://journals.sagepub.com/doi/abs/10.3102/003465430298487>

Ramdass, D., & Zimmerman, B. J. (2008). Effects of self-correction strategy training on middle school students' self-efficacy, self-evaluation, and mathematics division learning. *Journal of Advanced Academics*, 20(1), 18-41.

Ross, J. A., Rolheiser, C. and Hogaboam-Gray, A. 1999. Effects of self-evaluation training on narrative writing. *Assessing Writing*, 6: 107–132.

Wiggins, G. (1990). *Workshop on authentic assessment*. Unpublished materials. Rochester, NY: Consultants on Learning, Assessment, and School Structure. As seen in: Goodrich, H. (1996). *Student self-assessment: At the intersection of metacognition and authentic assessment*. Unpublished doctoral dissertation. Cambridge, MA.: Harvard University. 16-17. Available: https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Goodrich%2C+H.+%281996%29.+Student+self-assessment%3A+At+the+intersection+of+metacognition+and+authentic+assessment.&btnG=