

Journal of Trainee Teacher Education Research

Confidence vs Correctness: a study of how assessment for learning (AFL) impacts confidence and attainment in a Year 10 chemistry class studying Collision Theory

Kate Rose Charnley

(PGCE Secondary Chemistry, 2019-2020)

email: kr.charnley@gmail.com

Abstract

The interrelationship between secondary school student confidence and attainment in science is a phenomenon that has been well documented by the science education community. The purpose of this action research intervention was to acquire an in depth knowledge of the relationship between self-confidence and academic attainment, and to explore the impact that a confidence based assessment for learning (AFL) could have on student self-confidence and attainment in science. The study was implemented during a five lesson sequence on Collision Theory with a representative top-set Year 10 class. A review of the literature indicated that a significant correlation exists between assessment, student self-confidence and attainment. The analysis of the results of the impact of confidence based AFL showed that the confidence of each of the students in the science classroom is significantly increased, yet altogether each student's attainment remains similar. Further analysis using a questionnaire suggested that students with the lowest confidence levels in science could in fact experience the greatest increase in their self-confidence in assessments through the use of regular confidence based AFL.

Confidence vs Correctness: a study of how assessment for learning (AFL) impacts confidence and attainment in a Year 10 chemistry class studying Collision Theory

Kate Rose Charnley

Introduction

Educational researchers have long recognised that student self-confidence about their academic capabilities is crucial for their actual academic attainment and future education. Secondary school educators and administrators within the United Kingdom (UK) have a responsibility to bridge the transition of students from compulsory science education into higher education or vocational settings through the use of avant-garde strategies. Compulsory assessments in science, such as the General Certificate of Secondary Education (GCSE) within secondary schools, act as the basic requirements when it comes to building any career, but unfortunately appear at a truly vulnerable time in adolescence in terms of developing self-confidence. The knowledge of self-confidence studies versus assessment attainment has been explored successfully within higher education, however, it is rare to find it widely explored within schools. As a result, if such a goal in science education is to not merely perform compulsory GCSE examinations but in fact to encourage more students to build a career in Science Technology Engineering and Mathematics (STEM) and tackle the current STEMs skills shortage, then it becomes especially important for the science education community to research and understand in greater detail the impact such assessment for learning strategies could play on student self-confidence and attainment at the secondary school level, and how it could be improved.

As part of the Postgraduate Certificate in Education (PGCE) teacher training course at the University of Cambridge, Faculty of Education, this study was driven by a professional concern for the lack of self-confidence in science at a nearby high attaining secondary school. This action research intervention examining the impact that assessment for learning (AFL) could have on student attainment and confidence in secondary school science also aimed to improve my professional practice and possibly provide adaptable information for the future development of other teaching professionals.

Literature Review

A brief literature review of the relationship of student self-confidence and attainment within secondary school education and confidence based AFL is discussed in this section.

The problem

In my professional placement school, students were performing well on standardised test scores such as, double or triple GCSE science results, with the average scores higher than the local authority average and the England average for the class of 2018 and 2019. However, on closer inspection based on observations, teacher perceptions and student views, the self-confidence of students in science at the school appeared to be very low.

In a study performed by Andersen and Cross (2014), it was suggested that individuals with higher expectancy (defined as one's confidence to perform a task) for success in science are often more likely to study it further, a conclusion highly relevant to the problem stated above. In this empirical study, examining 19 259 ninth grade students from 944 schools (~ 27 students per school) from ten different USA states, it was found that 28 per cent of students of high ability were classed with low expectancy. This means that many of the students who usually performed highly in academic assessments and obtained good results did not in fact feel confident in performing these tasks. With such a large number of secondary school students being examined in this study and across a wide spread of geographical locations in America by Anderson and Cross, it construed very few impact limitations or lack of confidence in its results. The problem I have identified for this study is in strong agreement with these conclusions drawn from Anderson and Cross (2014) and as a result, it could be suggested that in general, the percentile of self-confidence in science within students aged 14-15 could in fact be very low for many high attaining students. Thus, if the self-confidence of high attaining students at the secondary school level was better understood it could be used to target specific aspects of pedagogies, like AFL in this study, that may promote a better relationship between these students' self-confidence and their attainment in science.

Self-confidence and attainment

This relationship between student academic attainment and the affective factor of self-confidence has been well explored and recognised by educational psychologists. Self-confidence has become one of

the most striking facets that can lead to enhanced academic performance in assessments (Bandura & Zimmerman, 1994; Foster, 2016; Gardner-Medwin, 1995; Komarraju & Nadler, 2013; Reed, Kirshner & Jolles, 2015).

Self-confidence is an ontological judgement defined as the students' belief in their ability to accomplish a given task or objective and perform well in school (Bandura & Zimmerman, 1994; Komarraju & Nadler, 2013; Wiliam, Lee, Harrison & Black, 2004). Many educational researchers have suggested that self-confidence is based strongly on a student's self-efficacy, an individual belief of their capacity to execute behaviours that are needed to produce specific performances in attainment (Bandura, 1997; Bong, Cho, Ahn, & Kim, 2012; Reed et al., 2015). Albert Bandura was the first to truly analyse students' self-efficacy as a factor that could inhibit or aid assessment performance. The extensive peer reviews and continual development observed throughout Bandura's work on self-efficacy provides confidence in the theory's validity. However, Albert Bandura's definition of self-confidence has one limitation in the context of this study (Bandura, 1997). Bandura suggested that self-confidence involves a student's anticipatory judgement in their chances of successfully answering a question correctly. Conversely, in this study, the self-confidence of students is taken in spite of the outcome that preceded by requiring them to accredit a confidence score at the time of answering the question. The method chosen in this study possibly helps the students to analyse their self-confidence more accurately in relation to the questions asked, therefore presenting a more informative method of determining the students' actual relationship between self-confidence and attainment in science. Supporting this, in a study performed by Bong et al. (2012) on a total of 746 elementary and middle school students from middle-classed public high schools in Korea - a location not matching this study performed in the UK thus resulting in a lack of impact limitation because it focuses on a different regional focus where the education received may greatly differ - the students' self-confidence was measured post question answering, an implementation that is directly relevant to this study, and found that self-confidence in itself was a superior forecast of academic attainment than other factors such as, self-esteem (a person's overall sense of personal value). As a result, it is correct to focus on the relationship of self-confidence and attainment in this study rather than another affective factor.

The cause of the problem

Among many academic factors that can influence student self-confidence and attainment, the transition into secondary school with its compulsory assessments is a widely recognised challenge by all education staff for its strong impact on self-confidence. Reed et al. (2015), proposed that previously easily high attaining students in primary school may experience dampened self-confidence on transitioning to secondary school due to having unrealistic expectations of the amount of effort required for secondary school assessments. Bandura (1997) listed adolescence itself as a difficult stage in development, with stress, tension, anxiety, and negative mood diminishing self-efficacy, and energy, enthusiasm, and optimism intensifying it. Bandura (1997), furthermore mentioned that even the combination of cultural stereotypes around science that exist in the world today can be easily intensified in the secondary school level and become damaging to student self-confidence. The prevailing attitude in society is that boys perform better in STEM, while girls perform better in languages and arts, leading to confidence biases in gender within school (ibid.). Pressures to adhere to conventional gender roles intensifies in early adolescence and has consequent effects on the respective performances of boys and girls in the two subject areas. The research study by Anderson and Cross (2014) also found this relationship and even concluded that expectancy (the student's confidence to perform a task) was positively related to ability but even high ability students who internalise stereotypes can have reduced attainment in the STEM domain.

Recently, Komarraju and Nadler (2013) performed a research investigation into the self-efficacy within undergraduates in the states. The study took place with 407 students ranging from seniors to freshman, from a broad ethnicity, of approximately equal gender, with a median age of twenty years, all studying a wide range of subjects including 11.8 per cent science. Such a broad integrated cross-curricula study, although based on a different age group of students, inspires confidence in there being few study design limitations and thus inaccurate results within this study. The findings of the study found that college Grade Point Average (GPA) positively correlated strongly to self-confidence, and negatively with test anxiety. The general consensus among social scientists that the most noticeable impact on student self-confidence and attainment stems from this test anxiety, creating negative affective states around assessment which will negatively influence student academic self-confidence and lead to a decreased academic performance (Anderson & Cross, 2014; Bandura & Zimmerman, 1994; Bandura, 1997; Komarraju & Nadler, 2013). With anxiety about assessments becoming the most detrimental factor affecting student attainment, with such an issue only heightened as previously

mentioned with students struggling with adolescence and key 'future determining' secondary school level assessments where the self-confidence in students studying science may already be at a low, only magnifies the importance of this study.

A possible solution: confidence based assessment for learning (AFL)

It is a truth universally acknowledged by educators that pupils within secondary school tend to react sensitively to assessment and that their self-confidence is highly vulnerable. When using AFL techniques, secondary school students' workings in a classroom (both oral and written) can range from a shot in the dark, a random guess with no confidence, to a certain answer of absolute self-confidence. Accompanying this, a teacher during these classroom assessment techniques would also not know which form the answers took. As a result, research has begun to show that confidence has in fact got a potential role to play within AFL (Bennett, 2011; Wiliam et al., 2004).

Regular integrated confidence based AFL is a method that could provide students with consistent exposure to regular assessment in order to try to reduce the anxiety behind tests and also provide students and teachers with an accurate perception of the student's own self-confidence in science. Confidence however, is difficult to assess reliably because students are always unpredictable and may be inclined to exaggerate their confidence in order to win approval from peers and teachers, or even vice versa to avoid attention if under confident. Many strategies, such as traffic lighting (RAG), hands up/hands down and think pair share etcetera, suffer from this problem. As such, a good confidence-based AFL is a type of assessment strategy that tries to accurately measure a student's confidence and attainment in one, without external influences effecting it. Typically, confidence based assessment is performed by asking students to give a confidence score for each of their responses to questions. If their response is correct, the number of marks that they receive for that question is the same as their chosen confidence level, however, if incorrect, negative marking takes place and a penalty of the confidence score is assigned. A study by Gardner-Medwin (1995), based on 77 male and 30 female students studying physiology in 121 computer based tests over a period of seven weeks covering 376 binary questions, showed that confidence based assessment for learning attempts to increase appropriate levels of self-confidence and supports realistic self-awareness in order to facilitate improvements to students' future attainment. On consulting students' perspectives, Gardner-Medwin (1995) found that students' informal reactions indicated that they appreciated confidence assessment, with only 12 per cent of students opting to not use confidence assessment

again. A limitation faced by Gardner-Medwin's (1995) empirical research was in fact that students may have found it difficult to accurately correlate their perceived self-confidence with a confidence score as no meaning was provided to the confidence levels they were expected to mark down. Bandura (2006), stated that for efficacy scales, "preliminary instructions should establish the appropriate mindset that participants should have when rating the strength of belief in their personal capability" (p.312). However, his suggestion of using a larger scale from 0 – 100 because they are more sensitive to students avoiding the extreme positions, would not be feasible for this study with secondary school students as it may become overwhelming for students of this age. Due to the age of the paper written by Gardner-Medwin (1995) a more recent study by Foster (2016) was reviewed. Foster (2016) analysed confidence based classroom assessment on secondary school mathematics students, using a simple scalar system of 0 – 10, and found that this pedagogical strategy in fact not only showed that confidence based AFL improved student self-confidence and attainment, it also encouraged students to focus on understanding the questions rather than just on their performance. The confidence based assessment was shown to influence students to think and behave rationally before giving an answer during the assessment, with students sometimes changing their answer in response to a confidence score in order to try to maximise their total score. The students gained a greater level of autonomy with the assessment.

Despite the positivity around confidence based assessment in science and mathematics by Gardner-Medwin (1995) and Foster (2016) respectively, implementing confidence based AFL in schools has its limitations such as: students experiencing anxiety about the assessment influencing their confidence in the subject; or because it could be viewed as a meaningless strategy used solely for the benefit of the teacher; or even because of the fact that many include negative marking in order to penalise guessing. Such limitations can be complicated and unfamiliar to school students, with the possibility of it even being viewed as a punishment for students based on what they do not know. Despite this, Wiliam et al. (2004), strongly suggested that AFL, if implemented properly by teachers, has been consistently shown to greatly influence student self-confidence and attainment, in particular among lower confidence.

Considering the evidence presented it can be seen there is potential for increasing student self-confidence and attainment within science lessons through the increase use of confidence based AFL. AFL strategies, in particular confidence based AFL, providing they are performed in the correct way,

can play a significant role in supporting and empowering student self-confidence and attainment within the science classroom.

Research Questions

This paper investigates the effect of the pedagogy, assessment for learning (AFL), on confidence and attainment. The following research questions are addressed in the context of a Year 10 class studying Collision Theory.

RQ1: Does the AFL impact student confidence in understanding Collision Theory?

RQ2: Does the AFL impact student attainment in understanding Collision Theory?

RQ3: How do students behold the AFL as an influence on academic self-confidence and attainment?

Outline of Teaching

This study took place in an 11-16 state school, rated “outstanding” by Ofsted, with approximately 1300 predominantly white and middle class students. The intervention of AFL was implemented in February and March 2020 during a series of five lessons which focused on Collision Theory, found in chapter six of the AQA GCSE Chemistry syllabus titled ‘Rates of Reaction’. Previous teaching of the Rates of Reaction lessons was performed by the usual class teacher, focusing on measuring rates of reaction, and by myself, focusing on the two GCSE required practicals for ‘Rates of Reaction’. All 32 students, 12 male and 20 female, of the Year 10 were invited to participate in the study. Four students were classified as pupil premium (PP), 12 as gifted and talented, and two students classified as having emotional or behavioural difficulties but did not require a teaching assistant. The class had three science lessons a week, 60 minutes each, with two different teachers in two different classrooms. Fifteen minutes of each of the five sessions taught by myself was devoted to collecting data for this intervention.

Table 1 lists the learning objectives and the type of data collected for each of the five lessons of the intervention taught by myself used for exploring the impact of AFL on student self-confidence and attainment on a top set Chemistry class. All 32 students were working towards sitting the AQA GCSE Chemistry Higher paper. The numbers of students for each predicted GCSE grade - Making The

Grade (MTG) target scores - are shown below in Figure 1. The mean MTG attainment is 7 and the modal grade is 8.

Lesson	Learning Objectives	Type of Data Collected
1	LO1: Explain that chemical reactions are dependent on collisions between particles. LO2: Define what is meant by a successful collision. LO3: Explain why not all collisions are successful.	Previous Knowledge Probe (PKP) Confidence Based Formative Assessment
2	LO4: Explain why a higher temperature produces a higher rate of reaction. LO5: Describe the effect of temperature on rate of reaction.	Confidence Based Formative Assessment
3	LO6: Explain what is meant by the term concentration. LO7: Describe and explain the effect on rate of increasing or decreasing concentration.	Confidence Based Formative Assessment
4	LO8: Explain what is meant by the term pressure. LO9: Describe and explain the effect on rate of increasing or decreasing pressure.	Confidence Based Formative Assessment
5	Summative Assessment – End of Topic Test	Confidence Based Questionnaire

Table 1: Learning objectives and type of data collected for each of the five lessons intervention

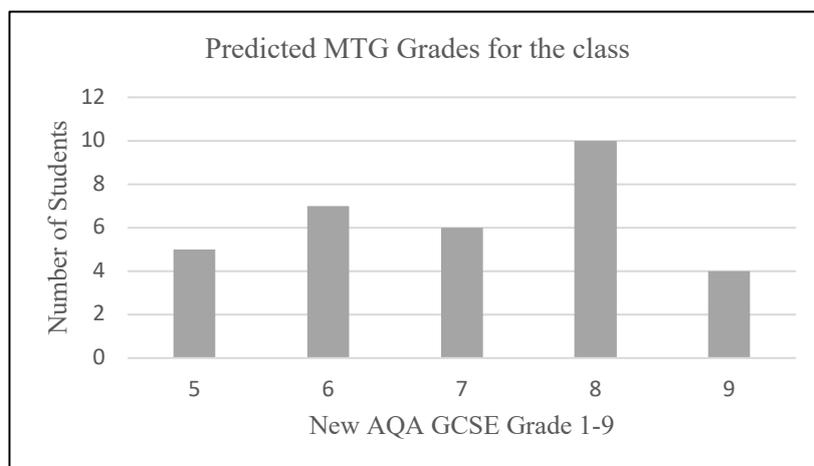


Figure 1: MTG target scores for predicted grades for the Year 10 class.

Methodology

It is firstly important to note that the class being studied has not been put together randomly as the classes were set by attainment. A mixed-methodology (qualitative and quantitative) action research study was chosen to explore the relationship of how confidence based AFL impacts the academic self-confidence and attainment in this Year 10 class. A mixed-methodology approach is suitable for this study because it combines the depictive and experimental essence of qualitative approaches with

the statistical foundations of a quantitative approach within an intervention. Furthermore, it makes it possible to quantify correlations and understand the factors of confidence based AFL that may result in improved self-confidence and attainment and those factors that can adversely affect it (Creswell & Creswell, 2018; Denscombe, 2010; Wilson, 2013a).

Action Research

Action research projects are driven by a professional concern to change or improve some aspect of teaching in the classroom, and to suggest ways in which that practice can be improved. According to Wilson (2013b) and Denscombe (2010), action research is a continuous process of planning, action, and reflection carried out by a teacher-researcher. This cycle can form a spiral if repeated, in order to continually refine approaches and practices to school-based problems, but it is important to note that in this action research study only one cycle was performed due to time constraints. Wilson (2013c) goes on to explain that due to the dual role of teacher-researcher there is significant knowledge and perspective on the given issue that an outsider would lack due to the proximity of the researcher to the project. In addition, Wilson (2013c) argues that, although potentially limited to a local scale, this research may be more impactful than research carried out by professional researchers, typically in universities, due to the high relevance of the practitioner and school. However, the same closeness can also be a limitation of action research as it may limit objectivity when it comes to analysis. Figure 2 represents an action research model involving six stages with one following the other proceeding cyclically in a continual loop, from identifying an issue or action to employ to refining and repeating the action research cycle.

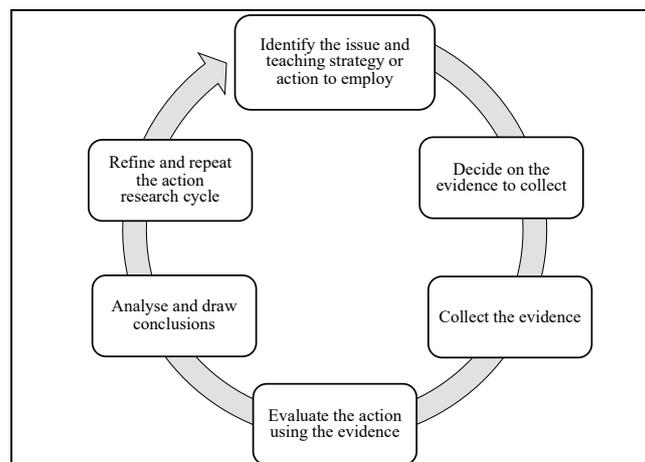


Figure 2: Action research model: a cyclical process

The aim of an action research intervention is not to end up with a universal truth about the best way to address the problem but a workable solution to one's professional issue that can also act as a set of suggestions that other practitioners may reflect on, criticise or enact.

Type of Data Collected and Analysis Performed

The majority of data collected during this action research intervention focuses on determining students' self-confidence and their attainment before and after the sequence of lessons on Collision Theory. The types of data collected are described below, in approximate chronological order.

Pre-Intervention Probe (PIP)

Prior to the intervention of confidence based AFL, baseline data was collected in the form of a PIP, whereby students wrote in their class books, using the confidence scale (as seen in Table 2), a current confidence score for science. This was used to assess students' current self-confidence in science prior to the intervention. Alongside this, I also collected the students' attainment levels from the previous summative assessment and their predicted GCSE grades in the form of their 'Meeting The Grade (MTG)' scores. The PIP was also used to try to assess existing knowledge of Collision Theory and also identify any common misconceptions based solely from an introductory question, "How would you define Collision Theory? And what factors might affect it?". All responses to the PIP, the confidence score, previous summative attainment, MTG scores and answers to the open-ended introductory question, were collected from students' class books and analysed.

Confidence Based Formative Assessment

Confidence based formative assessment is used to express confidence. In this study a 'Confidence Game' was played for an AFL strategy, and implemented for 15 minutes as a plenary activity for each of the four teaching lessons in the intervention sequence. A five question quiz on Collision Theory based on each of the learning objectives within that lesson was designed for each teaching lesson and used in order to ascertain students' responses and self-confidence level on the material of Collision Theory. Students were asked to ascribe for each question, a response and a confidence scale value which corresponds to how confident they felt the written answer was correct. The confidence scale (see Table 2) was to be given as a whole number from 1 to 5 and, for ease of the assessment implementation, each question was designed to act as a one mark question. As suggested by Bandura

(2006), to avoid complications a guide was given to students as to what these numbers for confidence could correlate to as shown below.

Likert Confidence Scale	1	3	5
Confidence Scale Guideline	Not Confident Unsure of answer	Some Confidence 50% correct	100% confident Answer is correct

Table 2: Guideline definition of confidence for each point on the Likert scale from 1-5

It was explained that a total mark for each of the AFL activities would be determined from calculating the sum of the confidence score values for the correct responses minus the sum of the confidence score values for the incorrect responses. The maximum score that could be achieved during each time the ‘Confidence Game’ AFL was played was 25, as this was the total score that could be attained had the confidence score been 5 each time and all 5 questions answered correctly, and vice versa, the lowest score that could be obtained was -25, as this was the total score had all the confidence scores been 5 each time but all 5 questions were incorrect.

The correlation between attainment and confidence for each question and each student was analysed using a scatter graph to determine how well the students could determine if their answer was correct for each question. The points on the scatter graph represent the mean mark obtained in each question by the whole class. The standard deviation bars show the variation of the student’s marks. The scores from each of the four AFL activities were also combined for each student to produce a percentage total, which was then converted using the schools grading system into the 1-9 GCSE grades and compared to the classes predicted ‘Making The Grade’ (MTG) target scores and the previous end of topic test results. A statistical analysis using a paired t-test was performed to determine if there was a significant change between the previous end of topic test and the confidence based AFL data.

The design of the confidence based AFL template issued to students is presented in Appendix 1. The template included a space for students to write any open comments regarding the task, and any open task comments would then be analysed.

Summative Assessment

The end of topic test was used as indicative proof of attainment during the intervention on Collision Theory. A paired t-test was performed on the summative end of topic assessment data on ‘collision theory’ and the previous end of topic test on ‘measuring rates of reaction’ was performed, and then

readily compared to see if the implementation of confidence based AFL had an impact on student attainment.

Questionnaire

A questionnaire was designed to probe pupils' views of how the confidence based AFL impacted their self-confidence and attainment post intervention. Designed as an exit ticket from class, students were asked to leave a sentence describing how they felt about the 'Confidence Game' and if they thought it had had any impact on their confidence and attainment in science classroom and also write down in their class books another confidence score assigning how confident they felt in science post-intervention using the confidence scale again (as seen in Table 2). Students' responses to the post-teaching questionnaire were analysed and modal responses were identified and coded into common categories based on the relation to the impact the confidence based AFL had on self-confidence and attainment. The questionnaire was included as a quick and easy method to form an impression of the common perceptions students possess in the science class about the relationship of self-confidence and attainment. However, as mentioned by Chaplain and Warwick (2013a), there are a few limitations to questionnaires that must be noted such as: students may become disengaged, answer questions hurriedly rather than answering honestly and carefully, alter responses to try to match the answers to peers or to give what they thought were the 'correct' responses in order to 'please' the teacher depending on the relationship.

Focus Group Interview

A focus group interview was designed to act as a secondary data source for student perception of the intervention and to determine which AFL techniques impacted student confidence and attainment the most. Purposive sampling, a form of convenience sampling, was thought to be best suited for this action research study as it is not performed on a truly random set of participants. Purposive sampling allows for the intentional selection of participants from the class from whom it would be able to extract relevant information.

Focus groups were planned to occur in an informal, nonintimidating space, to not hinder the participation from the group. Open-ended questions were drafted to align with the problem, the impact AFL could have on students' academic self-confidence and attainment, and the findings summarised and triangulated with their PIP responses and summative assessment scores (Chaplain &

Warwick, 2013b). Two of the main potential limitations of focus groups in this action research setting with a teacher-researcher could be that students are reluctant to share ideas in front of peers for fear of embarrassment or judgement, or that the pre-established relationships that exist between the researcher-teacher and the students could cause altered responses depending on the relationship (Wilson, 2013).

Triangulation of Data

Finally, Denscombe (2007) suggested that collecting data from a myriad of sources allows for data triangulation, thus enhancing the rigor of the study. Triangulation of data was used in this study within the results section, by using of multiple sources of data to examine the same research question. This naturally improves the accuracy and helps to gain a more complete picture of what the data gathered suggests. The planned data sources to address each research question are presented in Table 3 immediately below.

Overall Research Question	How does assessment for learning (AFL) impact student self-confidence and attainment in a Year 10 class studying Collision Theory?		
Data Source	Research Questions		
	RQ1 Does AFL impact student confidence in understanding Collision Theory?	RQ2 Does AFL impact student attainment in understanding Collision Theory?	RQ3 How do students perceive AFL as a way to improve the relationship of academic self-confidence and attainment?
PIP	✓	✓	
Confidence based AFL	✓	✓	✓
Summative Assessment		✓	✓
Questionnaire	✓	✓	✓
Focus Group	✓	✓	✓

Table 3: Planned data sources to address each research question

Validity and Reliability of Data Collected

Firstly, it is important to note that there is no control group present within this study as another class of the same type was not taught at the time of the intervention, therefore it is difficult to determine with certainty that the cause of these findings is in fact due to the implementation of confidence based

AFL. This makes it harder to demonstrate the reliability of the data collected and conclusions drawn. Despite this, although a true control is never possible in educational research because students are always unpredictable, a comparison of different classes could have been instructive for this intervention. The conclusions of this action research are therefore subjective, remaining tentative and contextualised, with a transparent guide about how all the data was planned and collected throughout the intervention, inclusive of explicit awareness of limitations for each data source.

Ethical Considerations

The action research intervention was conducted in accord with the British Educational Research Association (BERA) guidelines (BERA, 2020) and the Faculty of Education ethics form at University of Cambridge, which was completed and approved by the subject lecturer. In accordance with recommendations, all students, parents, teachers and head of department were informed of the intervention and those involved in the project given the opportunity to opt out or consent, inclusive of any interviews, surveys or audio recordings. Students were also assured that all data would be treated in line with the school's guidance on data protection and anonymised in the final report. The class followed the intended curriculum during the intervention and so were not disadvantaged compared to other classes. No decisions made were knowingly detrimental to the class progress but it is important to note that students' awareness may have caused different behaviours and attitudes to have been exhibited whilst the study was executed. All data collection was carried out within the scheduled lesson time.

Results

In this section the results of the action research intervention are laid out in chronological order of the data collected. It is also vital to say that due to unforeseen circumstances of COVID-19, Table 4 details the data that was actually collected before school closure. Focus group data across all three research questions could not be collected and also the summative assessment could not be undertaken for RQ3. Constraints on time due to school closures for COVID-19 also meant that more complete data collection (for example, the testing of more confidence based AFL strategies) was not feasible.

Overall Research Question	How does assessment for learning (AFL) impact student self-confidence and attainment in a Year 10 class studying Collision Theory?		
Data Source	Research Questions		
	RQ1 Does AFL impact student confidence in understanding Collision Theory?	RQ2 Does AFL impact student attainment in understanding Collision Theory?	RQ3 How do students perceive AFL as a way to improve the relationship of academic self-confidence and attainment?
PIP	✓	✓	
Confidence based AFL	✓	✓	✓
Summative Assessment		✓	
Questionnaire	✓	✓	✓

Table 4: Actual data sources collected to address each research question

Pre-Intervention Probe (PIP) Results

Baseline Confidence Data

Figure 3 presents a bar chart showing the baseline confidence data for the whole class collected from the pre-intervention probe (PIP), where students wrote down a confidence score using the confidence scale (as seen in Table 2) as a guide in order to assign their confidence in science pre-intervention. It shows that, excluding the students who did not place a baseline confidence score, the whole class’s baseline confidence scores pre-intervention ranged from 1, not confident at all, to 5, very confident in science, and have both a mean and median of 3 and a mode of 2.

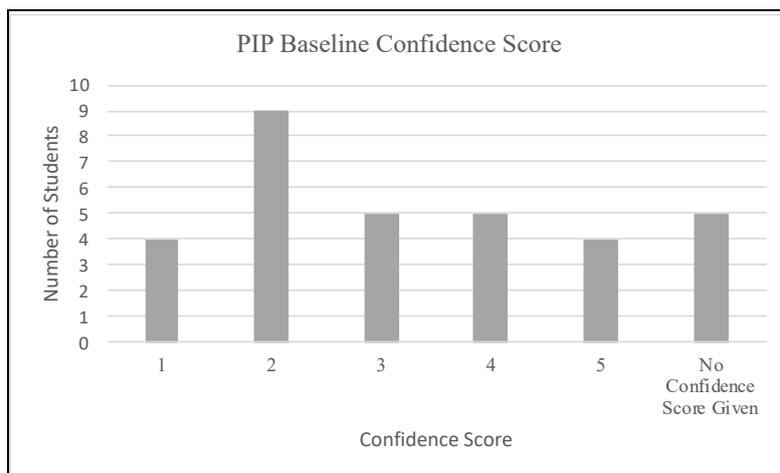


Figure 3: Baseline confidence scores collected from PIP for Collision Theory pre-intervention

The percentage of students who did not feel confident (56%), taken as those who selected ≤ 3 , was higher than those who selected that they did feel confident (28%). Indeed, when comparing these figures, it can be observed that twice as many students felt less confident than those who indicated they did. Interestingly, for various reasons, 16 per cent of students out of the class total did not write down a confidence score despite encouragement and as a result of this incomplete data were disregarded in the analysis of the baseline confidence data. However, if combined, approximately 63 per cent of the students in the class are shown to have very low self-confidence or were so underconfident that they did not even write down how they felt about their self-confidence in science.

Previous Attainment Data in Science

Figure 4 shows a bar chart presenting the whole class attainment in the previous end of topic test that occurred pre-intervention, based on ‘measuring rates of reaction’, compared to the whole class’s respective MTG target grades. The grades range from 5 to 9, in line with the New GCSE Grades, and it can be observed that there is a mean and median previous end of topic score of 7 and a mode of 8. The mean and median for the previous end of topic assessment attainment are the same but less than the mode, therefore the results show very little skewing. It can be observed in the bar chart that only slightly more students received lower grades than their predicted MTG grades. In particular, more students obtained a grade 5 and fewer students obtained a grade 9 in the previous end of topic test when compared to the MTG scores.

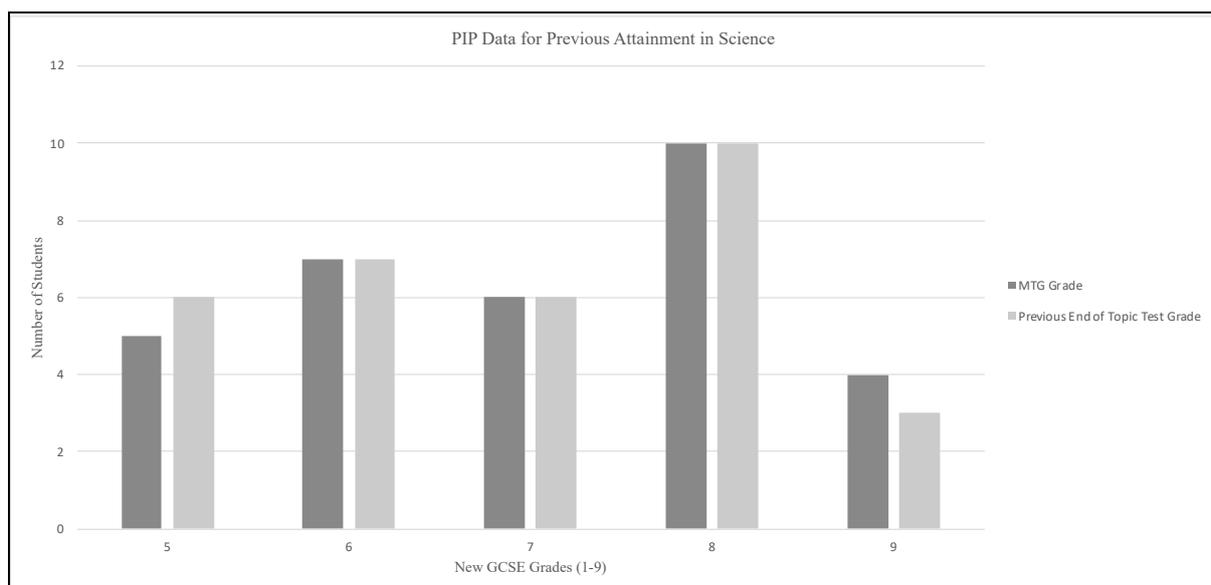


Figure 4: Attainment data collected from the PKP for the previous end of topic test

Pre-Intervention Misconceptions

The PIP model answers revealed very few misconceptions that exist around students’ pre-conception of Collision Theory. Firstly, the majority of the students defined Collision Theory correctly, deducing it easily as the theory of atoms colliding for a reaction. This is possibly because the definition of Collision Theory in books and educational websites is often written clearly and focuses on describing the key things that decide whether a particular collision will result in a reaction. In this case learners’ associative intuitive thinking of the words ‘Collision Theory’ were well established. A further factor that may have influenced the lack of misconceptions, was that the pre-questionnaire was given to students at the start of a lesson that followed on from an end of topic test based on earlier material in the ‘Rates of Reaction’ topic, measuring rates of reaction and GCSE Required Practicals. Therefore, students may have been answering the pre-intervention questionnaires with the recently covered pre-requisite knowledge of the topic predominately in mind, thus students’ initial self-confidence in science and towards Collision Theory may have affected their response to the PIP questions.

Confidence Based AFL

Figure 5 shows a bar chart that presents the mean mark obtained for the class in the confidence based AFL, with the four times that the Confidence Game was played shown on the same graph.

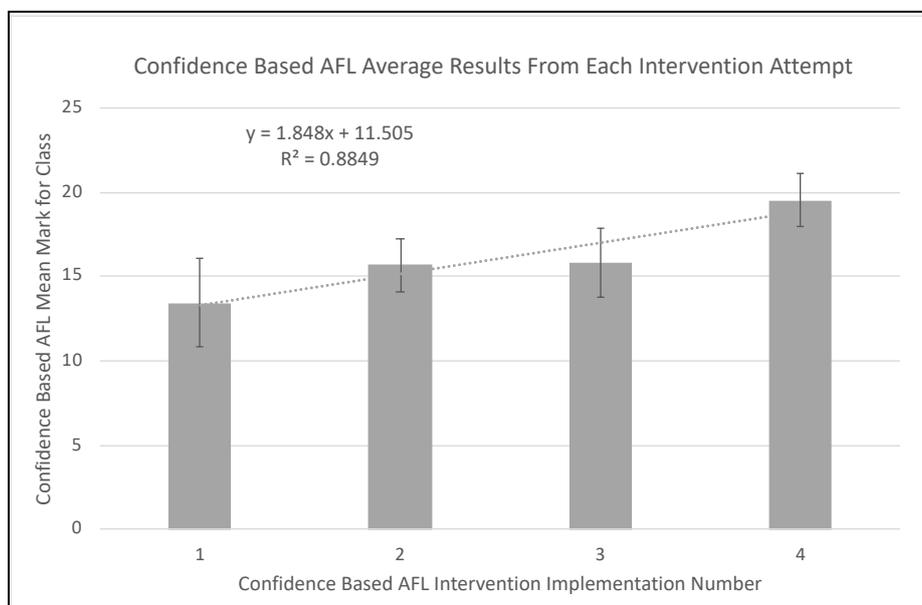


Figure 5: Confidence based AFL data, showing the average mark obtained from all 32 students for each of the four times the game was implemented as a plenary

There were five questions per AFL activity and an average of all the students across all five questions was taken to make it fair as each of the five questions could have been performed in any order. A medium positive correlation of 0.885 increase on the average attainment in the AFL task was found throughout the implementation. This could suggest that it is possible for the students' perceptions of themselves in terms of self-confidence to increase with continual use of confidence based AFL, and as a result an increase in attainment can be observed as indicated by their confidence based AFL scores on the graph. Overall it can be seen in Figure 5, that students' correctness and therefore attainment in the confidence based AFL increased by approximately 52 per cent post-intervention.

Figure 6 shows a line graph consisting of the mean mark of the class for each of the confidence based AFL questions including error bars of the standard deviation of the results for each question in turn.

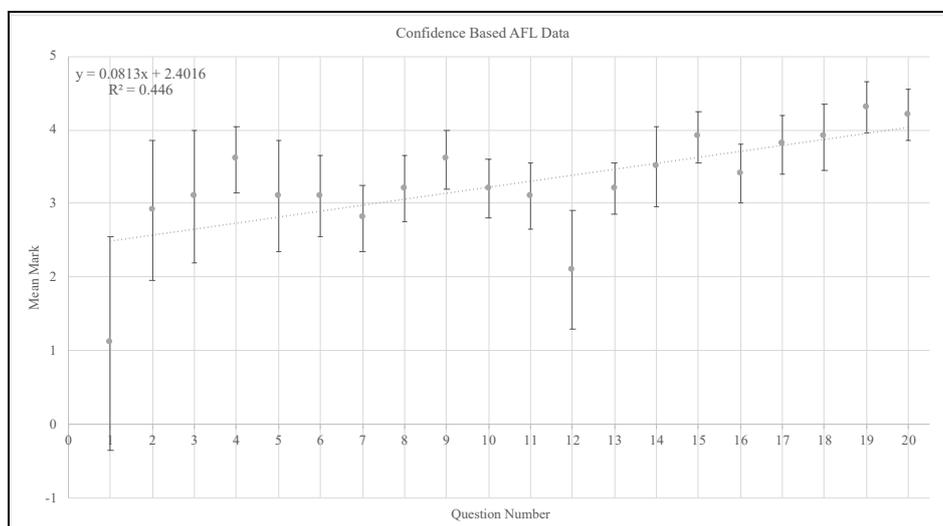


Figure 6: Confidence based AFL data, showing mean mark obtained from all 32 students for each question including the standard deviation

Looking more closely, initially the standard deviation on Figure 6 shows a broad range of attainment results, meaning that during the first questions many students struggled to correctly correlate their confidence with the correct score. Students with higher confidence scores at the start of the intervention may have viewed themselves as being more academically able in science and correspondingly students with lower confidence scores in the confidence based AFL may have considered themselves to be less able to answer the questions in the science assessments correctly. This could correspond to the varied AFL scoring on the early questions, shown by the largest standard deviation markers. Observing the standard deviation markers for each question they can be seen to

be getting smaller each time the confidence based AFL is performed. This suggests that individually students changed their perceptions of their ability in writing down the correct answer and also believed that the answer was correct thus improving their self-confidence. By the last confidence based AFL, the standard deviation for the answers is small, showing that the students were now better tuned to the relationship of their actual attainment and self-confidence in science. Interestingly, Question 12 in the confidence based AFL activity three, appears to be an anomaly. This could be due to the fact that the question appeared to be much more difficult than the previous questions, or that it was a question based on material that may have been less well explained by the teacher in the classroom.

Figure 7 presents a bar chart showing the number of students that have a certain predicted MTG GCSE grade and the number of students that obtained a certain GCSE grade in the AFL intervention. Note that this chart, and also Figures 8 and 9 later, displays the numbers of students in the class obtaining each grade rather than showing individual student performance.

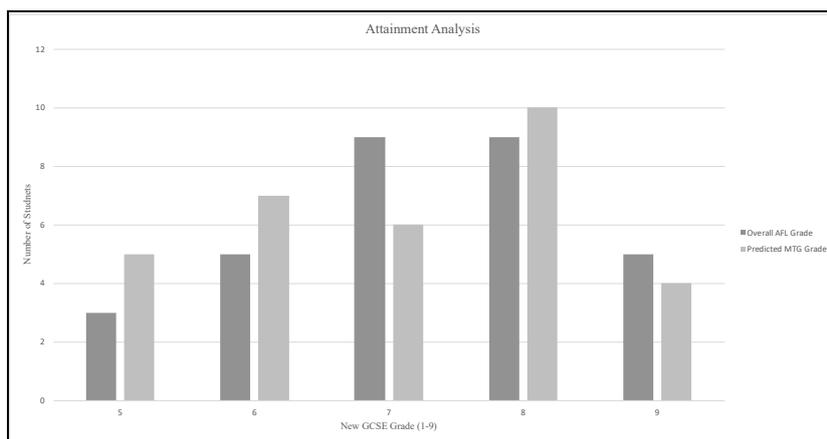


Figure 7: Comparison of students’ grades obtained overall in the confidence based AFL to their predicted MTG scores in science.

Further analysis on students’ actual attainment in the confidence based AFL was performed by adding up each student’s result in each AFL activity into an overall result out of a total of 100, as this was the maximum mark that could be obtained if for each of the four times the ‘Confidence Game’ was played the students obtained a maximum mark of 25. The results were then easily recorded as a percentage and converted into the GCSE grades using the schools ‘percentage-GCSE grade’ calculator used in the science department. Interestingly, the bar chart shows that students with lower predicted grades in fact scored more highly in the ‘Confidence Game’ resulting in a joint mode to be

observed at grade 7 and 8. However, on performing a paired-sample t-test, there was no significant difference in the grades obtained in the confidence based AFL (Mean (M) = 6, Standard deviation (SD) = 2.68, Sample Number (N) = 32) and the students' predicted MTG grades (M = 6, SD = 2.30, N = 32) in the conditions of Critical T-Value (t) = 2.00 and p-value (p) = 0.05. The paired t-test result was calculated to be zero, meaning that there was no statistical change found in the confidence based AFL attainment compared to students' usual grades obtained in assessments. Any observed differences could be due to chance or sampling error.

Summative Assessment Results

Figure 8 is a bar chart that presents the number of students who achieved each grade in the end of topic test, based on Collision Theory, compared to the whole class's MTG scores. It shows the whole class attainment in the end of topic test that occurred post-intervention on 'Collision Theory', ranging from 5 to 9 in line with the New GCSE Grades, and with both a mean and median of 7 and a mode of 8. The mean and median for the end of topic assessment attainment are the same but less than the mode, therefore the results show a little skewing to the left. The graph shows only slightly more students receiving higher grades than their predicted MTG grades. The biggest change was that more students in the lowest grade band increased their grade and more students obtained a grade 9 than were predicted (see Figure 8).

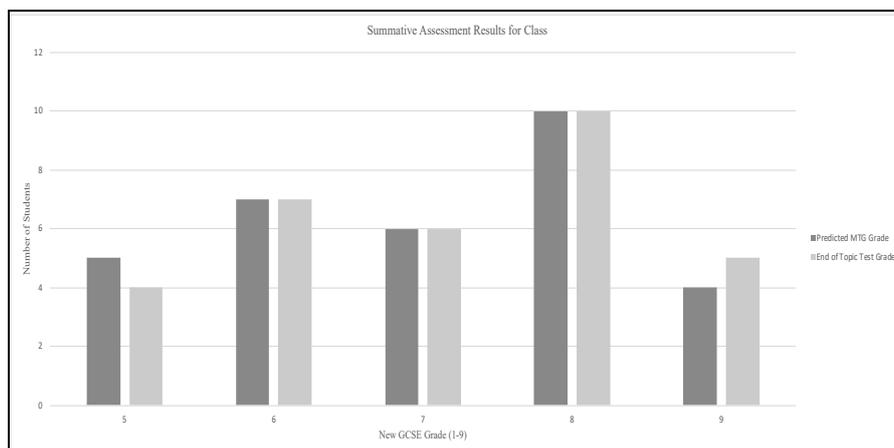


Figure 8: Attainment data collected from the end of topic test on Collision Theory

Figure 9 is a bar chart showing the number of students that obtained a certain GCSE grade on the end of topic test, based on Collision Theory, and the number of students that obtained a certain GCSE grade on the previous end of topic test, based on Rates of Reaction.

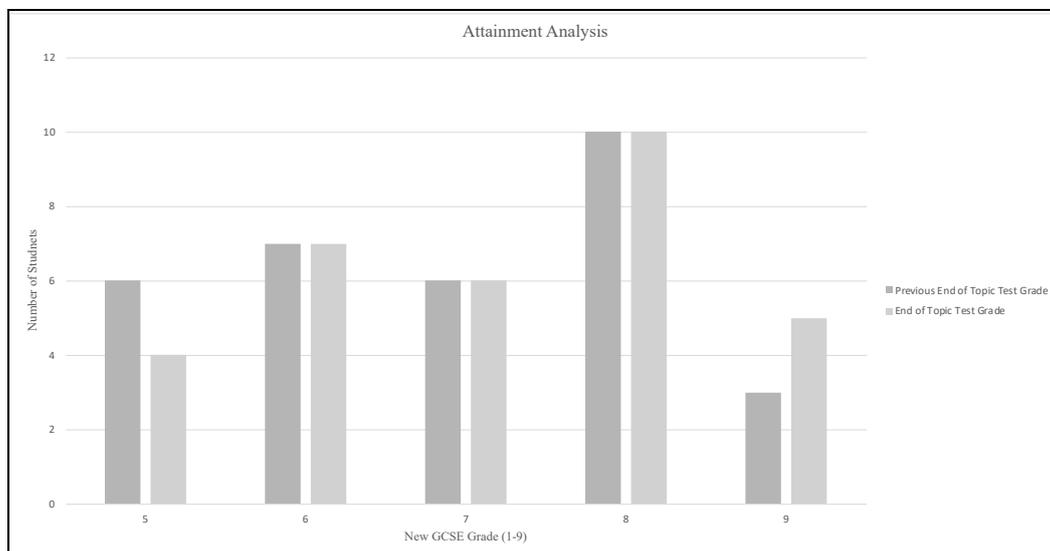


Figure 9: Comparison of previous end of topic test grades pre and post-intervention.

A paired-sample t-test was conducted to compare the results obtained from the previous end of topic test on ‘measuring rates of reaction’ performed pre-intervention, compared to the class results obtained from the end of topic test on ‘Collision Theory’. There was found to be no significant difference in the grades obtained in the previous end of topic test (Mean (M) = 5, Standard deviation (SD) = 2.31, Sample Number (N) = 32) and the grades obtained on the ‘Collision Theory’ end of topic test (M = 6, SD = 2.06, N = 32) in the conditions of Critical T-Value (t) = 2.00 and p-value (p) = 0.05. The paired t-test result was calculated to be lower than the critical t value and therefore the null hypothesis was accepted, and states that there was no statistical difference found between the pre- and post-intervention attainment.

Questionnaire

Comments Analysis

Comments were coded in more detail to result in producing three overarching themes, an overall view, negative views, and views that the AFL increased confidence and/or attainment, that were common throughout all students’ questionnaire responses. Figure 10 presents the three themes the

responses from the questionnaire are qualitatively coded into: an overall view, negative views and views that the AFL increased confidence and or/attainment. Figure 10 includes seven selected examples that were written by the students for each respective coded theme.

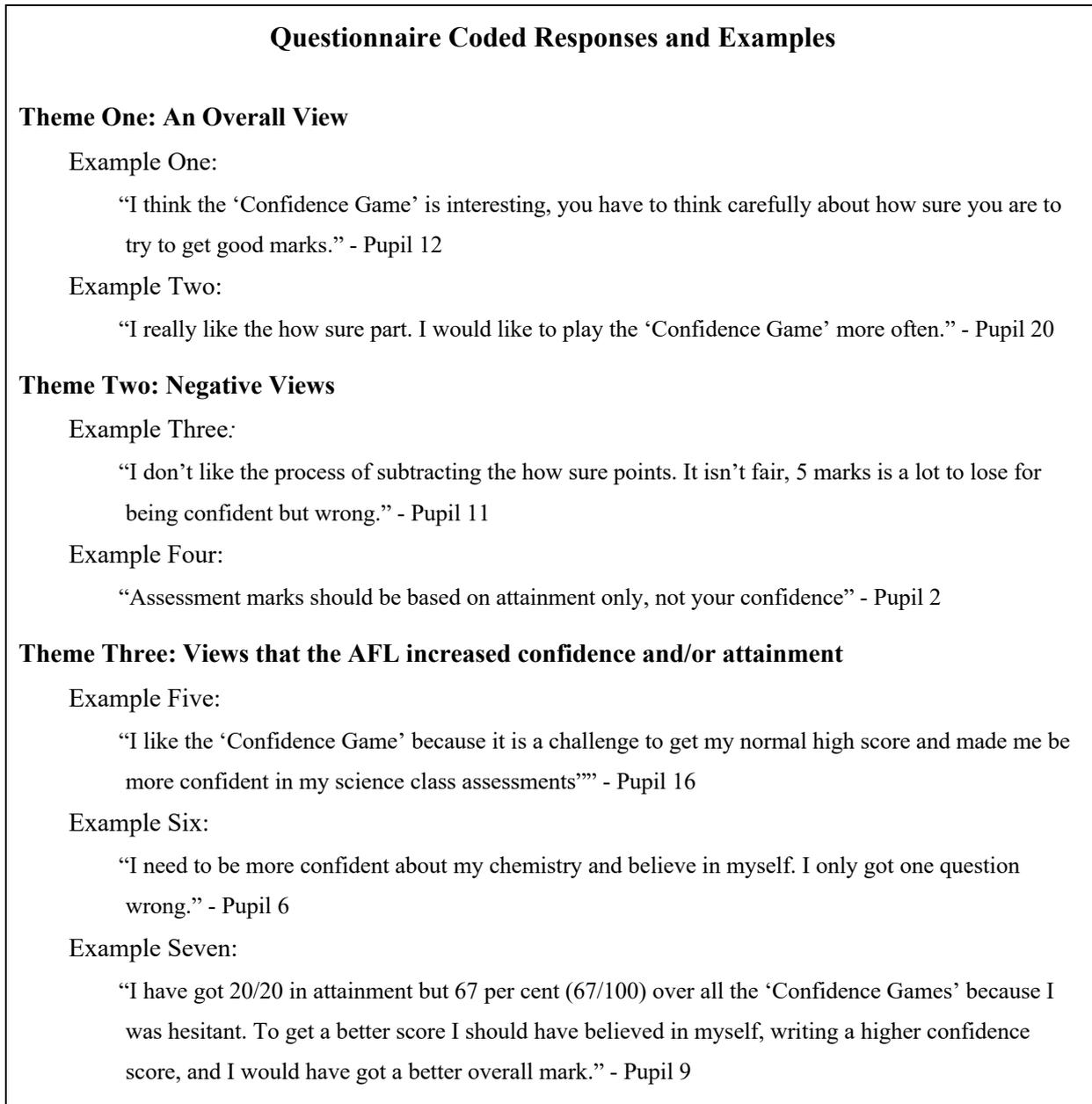


Figure 10: Questionnaire responses qualitatively coded with examples

Many of the responses written on the questionnaire by the students are consistent with expressing a clear overall opinion, or a positive or negative opinion about the confidence based AFL intervention. Of the 32 pupils who were present in class and answered the questionnaire, there were more positive responses about the ‘Confidence Game’ than detrimental responses. Despite there being no real

significant difference in student attainment over the course of the intervention, the foothold the confidence based AFL gave students upon identifying their correlation between self-confidence and attainment in science is vital. The results from the questionnaire show that students were not simply stating a preference for the use of the confidence based AFL simply because of the result on their success in assessments, but they were in fact surprised by the impact it could have. Many of the comments that were left on the exit tickets generally recorded a good enlightenment of how the confidence based AFL process worked and also the predicament of wanting to give a high confidence rating if the answer was correct but a low one if it might be incorrect. Many of these pupils also recorded that they had given more thought to their responses, taking more time and were careful in answering questions ensuring they focused on how sure they were. A very small number of pupils actually expressed problems regarding the use of the confidence based AFL because of the negative marking of the confidence scores which lowered their scores, findings that are also supported in Foster (2016) and Gardner-Medwin (1995).

Finally, the questionnaire found that a few students even directly stated or implied that the confidence based AFL, the 'Confidence Game', had made them more aware of the relationship between their self-confidence and attainment in science and in fact raised their self-confidence. These students addressed the problem in this intervention and believed that post-intervention they should believe more in their ability in regards to science assessments and could potentially obtain higher results because of this new found increased self-efficacy. A result similarly suggested in the study performed by Komarraju and Nadler (2013).

Confidence Score Analysis

Figure 11 is a bar chart presents the class's confidence scores post-intervention compared to the class's confidence scores pre-intervention, including the number of students who chose not to give a confidence score.

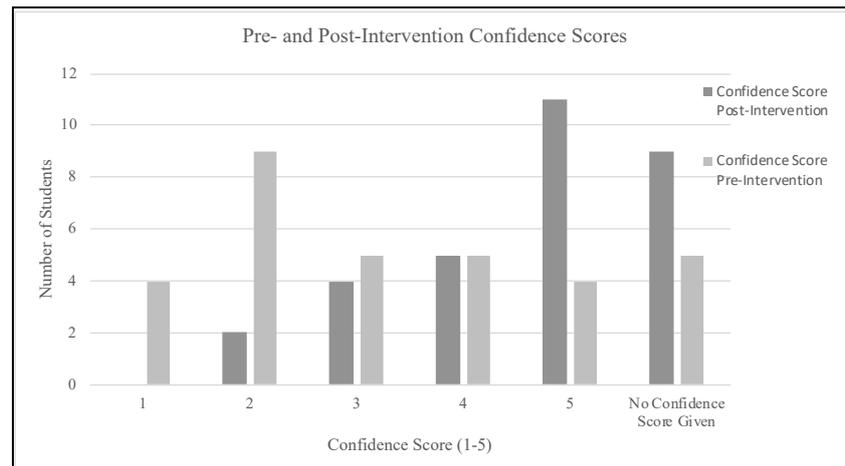


Figure 11: Pre- and post-Intervention confidence scores

Figure 11 shows what could only be described as a dramatic increase in students' overall confidence scores in science post-intervention, excluding the data where no confidence scores were given. A paired-sample t-test was conducted to compare the results of the baseline confidence scores in science obtained pre-intervention, compared to the class confidence scores obtained post-intervention. It was hypothesised that there would be no significant change between the pre- and post-intervention confidence scores. However, there was found to be a significant difference in the confidence scores of the class obtained in pre-intervention ($M = 2$, $SD = 1.8$, $N = 32$) and the confidence scores of the class obtained post-intervention ($M = 4$, $SD = 4.1$, $N = 32$) in the conditions of Critical T-Value (t) = 2.00 and p-value (p) = 0.05. The paired t-test result was calculated to be greater than the Critical T-Value and therefore the null hypothesis is rejected, and it confirms there was a statistical difference found between the pre- and post-intervention attainment.

An interesting note on the data obtained for confidence scoring, is the fact that there is an increase in the number of students absenting from writing down an overall confidence score in science. This may be due to the fact that students have rushed through the questionnaire to the end of class and the end of a school week, or the fact that they deemed it pointless, or the fact that the students generally did not feel confident enough to write down a confidence score. This issue may have been resolved during the focus group interview that could not be performed due to school closure, as the selection of a few over-confident students and underconfident students and their corresponding PIP, questionnaire and attainment data may have determined why students left the overall confidence score blank on the questionnaire.

Discussion

The results will be considered in triangulation under the headings of the three research questions given earlier: does AFL impact student self-confidence, does AFL impact student attainment, and how do the students perceive AFL to impact the relationship of confidence and attainment?

RQ1: How does assessment for learning affect student confidence?

It was anticipated from previous studies by Foster (2016) and Gardner-Medwin (1995), that with regular use of confidence based AFL students' calibration of self-confidence and attainment across Collision Theory would improve over time because students become better attuned to noticing their confidence levels when answering a question. As can be observed from Figure 11, student self-confidence did in fact increase over the use of confidence based AFL. Excluding the number of students who did not provide an answer, more students post intervention can be seen to have a much higher self-confidence score post implementation of the use of the confidence based AFL. In comparison to questionnaire results, many students' comments on the exit ticket recorded that they would have given different confidence scores now that they had become more confident in their relationship of attainment if given the opportunity to attempt more confidence based AFL. Thus suggesting that they had gained a better perception of their reliability in attainment with regards to Collision Theory.

It is important to remember that caution is needed within discussion because the students involved in the intervention were not a truly randomly selected sample, a limitation also seen in research performed by Foster (2016). The fact that the correlation, although statistically significant and high, is by no means perfect and should be seen as tentatively positive. Furthermore in-depth longitudinal studies which are fully randomised would be needed to determine whether this is the case.

As an extra observation of results, interestingly girls initially demonstrated much lower confidence than their attainment would justify whereas boys were deemed as over-confident in class. This observation is consistent with previous suggestions from Anderson and Cross (2014). However, further study is needed to determine to what extent the calibration of self-confidence and attainment is impacted by gender across other science topics.

RQ2: How does assessment for learning affect student attainment?

Figure 12 is a bar chart that presents the number of students that obtained a certain GCSE grade for the AFL intervention, predicted MTGs, the previous End of Topic Test and also the End of Topic Test. Overall no significant change was seen post-intervention compared to pre-intervention. The results, did not show a significant difference in attainment between students pre- and post-intervention attainment on analysis of the class end of topic tests or confidence based AFL data. This study unfortunately, does not demonstrate a statistically significant change in student attainment with the implementation of confidence based AFL, contrary to studies performed by Foster (2016) and Gardner-Medwin (1995). However, it is important to state that this study involved a short intervention, and had this study been as long or longer than Gardner-Medwin's study, a different outcome may have been observed.

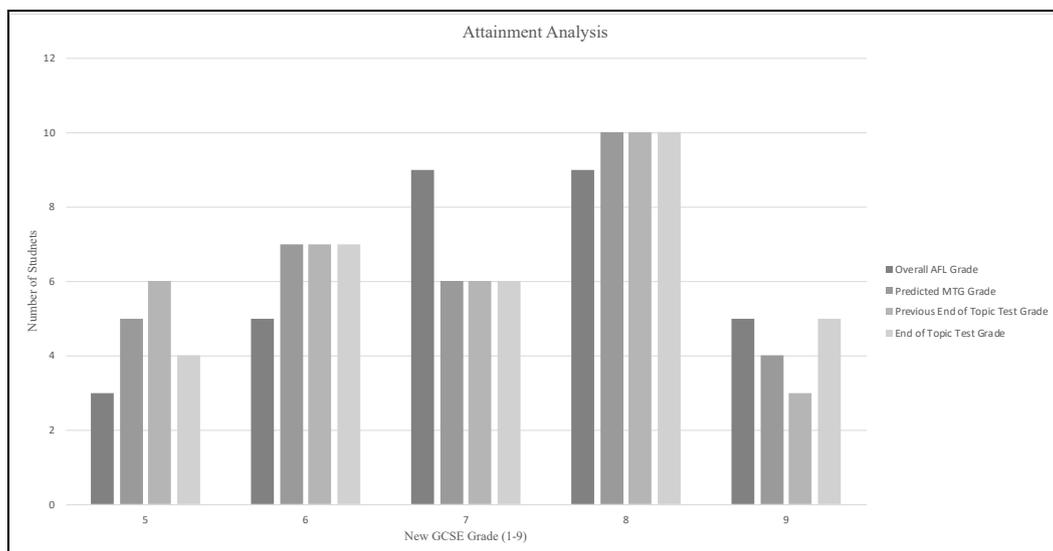


Figure 12: Attainment analysis comparing predicted MTG grade, previous end of topic assessment, end of topic assessment and the overall confidence base AFL attainment

It is possible that the intervention just in fact stabilised student attainment over the course of the intervention and prevented attainment from falling, rather than increasing, but it is not possible to conclude this from the collected data. Other than limited perceived increase in attainment from students, the intervention may have also likely been lost on attainment due to background influencing factors such as the change of a classroom teacher, the topics covered and the approaching end of term amongst others. Whilst this study did not prove the intervention was beneficial it also did not prove

it had unintended negative consequences. Considering the potential benefits and the ease of implementing this intervention a longitudinal study would be worthwhile.

Despite this, the perceived self-confidence within the class, observed by myself and the usual class teacher, was seen to increase in general through the use of the confidence based AFL. The observations of students showed that there was roughly three types of confidence versus attainment categories of students within this class: high confidence and lower attainment, low confidence and high attainment, and finally students who were naturally well correlated in these categories. Interestingly, it was observed that the confidence based AFL had the greatest impact on the low confidence students of high attainment, resulting in these students believing in their ability within science. This observation correlated well with these students' comments on the questionnaire, as seen in Example 6, Figure 12.

The manner of implementation of these activities may have had an influence on the amount of self-confidence observed. For example, 15 continuous minutes of the confidence based AFL for four lessons in sequence, may have resulted in students' attention to the task being enhanced and thus their independently gaining a deeper perception of their relationship of confidence in science compared to their attainment. It was not a small one-off task interspersed with other activities. This suggests that the more thoughtful students had considered the concepts within the task and questionnaire since completing the pre-intervention questionnaire and responded more carefully to task questions and more accurately to the post-intervention questionnaire. On the other hand, it is possible that some students may also have viewed repeating tasks such as the questionnaire, as pointless and as such did not complete them with due care and attention despite my best efforts to indicate the importance of doing so.

RQ3: How do students and teachers perceive AFL as a way to improve the relationship of academic self-confidence and attainment?

It is clear that pupils understood the confidence based AFL activity very readily. The observed medium to large positive correlation portrayed in Figure 5, implied that pupils were attempting to give true confidence levels, an ideal that is also reflected in the comments given from the questionnaire. It seems likely that this confidence based AFL approach could give genuine confidence levels of students due to the fact that it affects their attainment directly, and because it reduces students inflating their levels of confidence in order to make a positive impression on the teacher and also discourages guessing answers. Having more accurate information about the pupils' real

confidence levels could enable the teacher to intervene in more effective ways to help increase student confidence and attainment in science assessments. Zimmerman and Bandura (1994) suggested similarly, that if data such as this was collected over whole topics by teachers it could reveal specific areas where the whole class feels inefficacious or the specific material individual students are struggling with, and strategies for the deficient component could be addressed and exercises transcribed. However, some caution is needed here because the schools and students participating in this study were a convenience sample and may not be representative of schools and students more generally. Further study would be needed to determine teachers' perceptions of the usefulness of confidence based AFL on student confidence and attainment in science.

Conclusion

The study has revealed that student self-confidence is a vital determinant of academic attainment in assessments. Further research should take place over an extended period of time, such as an academic year, so comparisons to other longer running studies can be more easily made. The extended time frame would also allow for more topics to be covered that may result in different outcomes for students depending on the topics. Expanding the number and nature of classes involved may also produce different outcomes, it could be that lower attaining students or different year groups respond differently to this intervention. Possibly a few possible hypotheses are: that if the intervention was performed on lower attaining students could the confidence based AFL improve their academic results whilst focused on also potentially improving their confidence.

Many students are faced with diverse challenges around Secondary School, and as such teachers must work in collaboration to help boost student self-confidence during this crucial time of compulsory assessments. A lack of confidence or a fear of failure can be a huge handicap and it is the belief 'I can' not 'I can't' that alleviates stress and anxiety, opening the door to the assimilation of knowledge. This self-confidence can be enhanced, as shown in this study, through the use of confidence based AFL. The confidence responses can therefore potentially be of considerable value to teachers during formative assessments, since they offer a reliable way of probing students' true self-confidence in science. Furthermore, it could be recommended that teachers should follow students more closely through the simple tracking of results from AFL as it can be used to help to inform teachers of the progress of pupils. Tracking of AFL data could distinguish areas in which specific pupils may need more support, or even help to spot a problem within a whole class, allowing teachers to make more

informed choices regarding the planning of lessons and formation of strategies to boost future attainment.

References

- Andersen, L & Cross, T. L., (2014). Are students with high ability in math more motivated in math and science than other students? *Roeper Review*, 36 (4), pp.221-234.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Worth Publishers.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In A. Bandura, *Self-efficacy beliefs of adolescents* (pp.307-337). North Carolina: Information Age Publishing.
- Bandura, A & Zimmerman, B. J., (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal*, 31 (4), pp.845-862.
- Bennett, R. E. (2011). Formative Assessment: A critical review. *Assessment in Education: Principles, Policy and Practice*, 18 (1), pp.5-25.
- BERA. (2020, April 21). *Ethical guidelines for educational research*. Retrieved from British Educational Research Association: <https://www.bera.ac.uk/researchers-resources/publications/ethical-guidelines-for-educational-research-2018>)
- Bong, M. C., Cho, C., Ahn, H. S., & Kim, H.J., (2012). Comparison of self-beliefs for predicting student motivation and achievement. *Journal of Educational Research*, 105 (5), pp.336-352.
- Chaplain, R & Warwick, P., (2013a). Questionnaires - Research with younger children: issues and approaches. In E. Wilson (Ed.), *School based Research: A Guide for Education Students* (pp.71-73). LA: SAGE Publications.
- Chaplain, R & Warwick, P., (2013b). Focus Groups - Research with younger children: issues and approaches. In E. Wilson (Ed.), *School based Research: A Guide to Education Students* (pp.67-70). LA: SAGE Publications.
- Creswell, J. W. & Creswell, J. D., (2018). Preliminary Considerations: Mixed Methods Design. In J. W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Approaches* (pp.14-15). LA: SAGE Publications.

- Denscombe, M. (2010). Action Research. In M. Denscombe, *The Good Research Guide: For Small-Scale Social Research Projects* (pp.125-136). London: Open University Press.
- Denscombe, M. (2010). *The Good Research Guide: For Small-Scale Social Research Projects*. London: Open University Press.
- Foster, C. (2016). Confidence & Competence in Mathematical Procedures. *Educational Studies in Mathematics*, 91 (1), pp.271-288.
- Gardner-Medwin, A. R. (1995). Confidence assessment in the teaching of basic science. *Research in Learning Technology*, 3 (1), pp.80-85.
- Komarraju, M. N. & Nadler, D., (2013). Self-efficacy and academic achievement. Why do implicit beliefs, goals, and effort regulation matter? *Learning and Individual Differences*, 25 (1), pp.67-72.
- Reed, H. C., Kirschner, P. A. & Jolles, J., (2015). Self-beliefs mediate math performance between primary and lower secondary school: A large-scale longitudinal cohort study. *Frontline Learning Research*, 3 (1), pp.35-54.
- William, D., Clare, L., Harrison, C., Black, P., (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy and Practice*, 11 (1), pp.49-65.
- Wilson, E. (Ed.). (2013a). *School based Research: A Guide for Education Students*. LA: SAGE Publications.
- Wilson, E. (2013b). What is educational action research? In E. Wilson (Ed.), *School based Research: A guide for education students* (pp.231-255). LA: SAGE Publications.
- Wilson, E. (2013c). Research Design. In E. Wilson (Ed.), *School based Research: A guide to education students* (pp.79-89). LA: SAGE Publications.

Appendix 1

Confidence based AFL template student sheet

Confidence Based Assessment for Learning Task

For each of the five questions below, write a scientific answer in the box labelled ‘Answer’. For each question, also write in the box labelled ‘How Sure’ a whole number from 1 to 5 to indicate how confident you are that your answer is correct.

On this scale 1 is completely unsure and 5 is completely certain that your answer is correct, and a 3 would indicate that you believe you are 50 per cent correct. Acceptable numbers to write in the box are therefore: 1, 2, 3, 4 and 5.

The total mark for this activity will be calculated as the sum of the “how sure” values for the correct responses minus the sum of the “how sure” values for the incorrect responses.

The highest possible mark you can achieve is 25, this corresponds to a confidence score of 5 being written each time and every answer to be correct, and the lowest mark you can achieve is – 25, corresponding to a confidence score of 5 being written each time but every answer is incorrect.

<i>Number</i>	<i>Question</i>	<i>Answer</i>	<i>Leave Blank</i>	<i>How Sure?</i>
1				
2				
3				
4				
5				

Comments?

Total =
