The impact of teaching an inquiry-based scheme of work on pupils' attainment and critical thinking skills; and on the pupils' and teachers' perceptions of inquiry-based teaching in science.

Emma Atkinson

MA by Research

University of York

Education

June 2016

Abstract

This research was conducted to test the researcher's view that when pupils learn science through inquiry-based teaching strategies, they improve their understanding of the subject and develop critical thinking skills. Although there is much research into inquiry-based teaching and learning, definitions are varied, and as such this research developed its own on which to base the intervention. Research into critical thinking skills focuses on older adolescents and young people, and therefore assessing 12-13 year olds, as in this research, is infrequent.

Quantitative and qualitative data was collected using a quasi-experimental format. Data for pupils' attainment and critical thinking scores were collected both pre- and post-intervention, and this was analysed alongside perceptions of inquiry-based teaching stated in pupil and teacher focus groups.

The intervention was a three-topic scheme of work taught to half of the cohort with the other half being taught the same content but using their teachers' normal approach. The intervention was written by the researcher and was based on research into the strategies termed inquiry-based following the literature review.

It was found that there were no gains in attainment for pupils being taught using the intervention rather than the normal style of teaching compared to the control group. Gains in critical thinking were found in the treatment group but were not significant, and therefore, it was concluded the inquiry-based teaching did not have a positive effect on either pupils' understanding of science or their critical thinking skills.

Pupils' perceptions highlighted that they did not enjoy the group work, open-ended nature of inquiry-based learning, and missed the structure of creating a set of notes. Teachers believed that pupils' critical thinking skills would improve using inquiry-based techniques, but that pupils required more training in the skills needed to make this type of learning successful for this to take place.

Table of Contents

Acknowledgments

In both conducting my research and writing up this thesis I have encountered many challenges. Firstly, due to the scale of the research, I had to get my department on board at school which required a team effort and compliance in teaching-style which went well beyond what would normally be required. The support of the senior leadership was essential in signing off both the research itself and the ethical consent and was agreed without hesitation. The support staff in the school, such as IT and science technicians, were invaluable in ensuring both the assessment instruments and lesson intervention went to plan over a number of lessons. Finally, the pupils with the support of their parents were compliant and enthusiastic in both accepting a new style of teaching, completing extra assessments, and volunteering for focus groups.

In every stage of my research I have had to learn both about carrying out social science research and the language and structure needed to write it up, this whole process would have been far less successful without my mentor Jeremy Airey, from the University of York. His advice, manner and availability have allowed me to develop the skills and confidence required to complete this thesis.

I would like to thank all the people mentioned above as their support and willingness to adapt have made this project successful.

Author’s Declaration

All of the work contained within this thesis represents the original contribution of the author.

This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1- Introduction

1. Introduction

This research is inspired by the researcher's interest in the use of inquiry-based teaching strategies. After outlining the research questions in the next paragraph, the practitioner's perspective will apply a personal context and justification for this research piece. This will be followed, in the next section, by setting out the definitions of key terms being used. The research design and methods used to investigate each question will then be briefly detailed, and the chapter will then be concluded with a description of the following chapters.

1.1 Research Questions

Research will be divided into three research questions with a sub-question within the findings from questions 1 and 2, these are listed below.

1) What is the impact of an inquiry-based scheme of work on pupils' critical thinking skills?

a) Is there a delivery variable that influences these effects?

2) What is the impact of an inquiry-based scheme of work on pupils' attainment?

a) Is there a delivery variable that influences these effects?

3) What are pupils' and teachers' perceptions on the use of inquiry-based teaching strategies in science?

1.2 A Practitioner's Perspective: Research Background and Justification

There are a number of reasons why I am interested in this research topic, which I have refined to produce the question above. As a science teacher I am immersed in the business of trying to engage pupils in the science curriculum. In this privileged position I have developed a practice-based understanding of what enthuses pupils and also what helps them learn; these two outcomes can often become mutually exclusive. Often in schools emphasis is on short-term goals; the next assessment or report to parents, or proving learning has taken place in a lesson. All of these seem to lead to the development of low-order recall at the cost of higher cognitive development. I, like other teachers, want to ensure that pupils develop reasoning skills, resilience, evaluative capabilities and the ability to collaborate, but the performance-focussed education system seems to act against this.

In Science the role of practical work is an important one, especially it seems, in the UK. From years of practise and study of current research, I can conclude that this area of teaching can be handled very badly in terms of developing pupil understanding. The formulaic nature of practical work, followed by data collection, data analysis and finally hoping that the results will represent the scientific phenomenon, has long been flawed. There are examples of it being done very well but there is definitely room for improvement. Having used resources produced by CASE (Adey, 1999) and UPD8 (ASE) I have seen how, with a lot of time and creativity, the use of inquiry-based teaching strategies can bring together ideas, practical work, data and positively engage pupils in learning. The development of such lessons is time-consuming and complex. and as such they are not frequently used in classrooms.

One of the reasons that inquiry-based strategies seem less popular amongst teachers is that they do not give immediate feedback on pupils' 'understanding' or, more realistically, their ability to recall scientific information or answer specific lesson-related questions. The development of other skills that may come from inquiry are undervalued as they are not assessed, and so their importance is diminished. However, in my experience, and from research ascertaining other teachers' views, there is a consensus that developing skills such as critical thinking are seen to be as important to the education of a child as the recall of facts. These skills help them to achieve highly in the education system and in careers.

Therefore, for all the reasons discussed, the research title has been formed. Rather than taking a piecemeal approach to inquiry-teaching, as is the current practice in my school, I will write a scheme of work with a clear route through the programme of study. By producing a considered inquiry-based scheme, the value of this teaching strategy will be investigated in three areas; pupils' attainment, pupils' critical thinking skills and pupils' and teachers' perceptions of inquiry-based teaching.

As a practitioner, it is clear the impact that Hattie's (2008) Visible Learning has had on staff training and school interventions. As such, it seems a good place to start when reviewing the literature. As a meta-analysis it also generates a list of contributory research which became a starting point for the more in-depth analysis presented in the next chapter. The results that Hattie presents show that inquiry-based learning did not provide improvements in the 'zone of desired effects' (see Fig. 1.1). Hattie's definition of 'zone of desired effects' is based on pupils making better progress by a particular intervention (e.g. Inquiry-based instruction) than the average progress made by all well-implemented interventions or changes at a school or classroom level. It also reflects the expected attainment progression for pupils in their school career.

The results do not support my belief about the value of teaching using inquiry-based strategies but his conclusions provides a solid starting point for discussion in the next chapter.

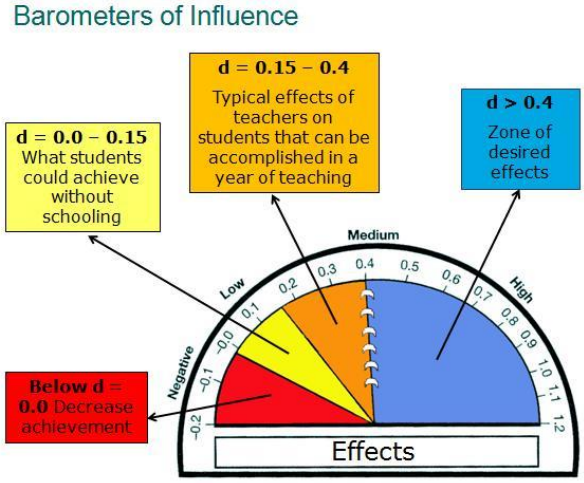


Fig 1.1 Barometer to show the effect of a particular intervention on pupils' progress (Hattie, 2008)

1.3 School Context

The research school is a larger than average sized urban 11-18 secondary school in a historic city in the North of England. Most students are of White British heritage. Fewer students than average are eligible for free school meals and the proportion of students with special educational needs and/or disabilities is below the national average. In this school, all young people follow the National Curriculum and at the age of the research cohort, in lower secondary science, the pupils study biology, chemistry and physics as a co-ordinated science. Pupils are taught in mixed ability groups, with three hour-long lessons per week. The science department is comprised of thirteen teaching staff with experience ranging from newly qualified to thirty years. All staff are science graduates holding a teaching certificate. The researcher is a teacher in this school and has the co-operation of the department and support of the senior leadership in carrying out this research.

1.4 Key Terms and Definitions

The following chapter which reviews the literature will set out in detail the challenges in defining some of the key terms. However, for ease of readership, below is a list of the main key terms and their definitions (*Table 1.1*), as they will be used in this research. It is particularly important that these are introduced at the start as there is not an agreed definition for many of them in existing research.

However, one term does need a more detailed explanation due to its importance in the research title, ‘inquiry-based teaching’. The literature surrounding this topic is vast, and has no agreed definition. Although there are ideas common to many, there is no clear consensus. The definition in this research is broad in order to allow the delivery of subject knowledge within an over-arching philosophy where many teaching strategies can be employed in order to meet the needs of individual learners. The definition also needed to produce a simple framework in order to create resources and lesson plans.

The guiding principle of inquiry-based teaching is that the pupils are using their natural curiosity, current schema and critical thinking abilities to develop and construct meaning, hypothetically leading to a long-term development of their cognitive structure and critical thinking skills.

Learning is gained from interacting with the scientific concepts being taught; either with primary resources, such as practical work and computer simulations, or secondary sources in the form of data and research information. Pupils must be investigating an authentic question, or the stages of the scientific inquiry-cycle (*Fig 1.2*) that follow on from one, this question can either be provided or negotiated. Pupils will follow elements of the cycle to investigate it using secondary skills such as research methods, discussion, questioning, negotiation, analysis, reflection and evaluation to determine their solution. Teaching techniques will focus on pupils being cognitively active with behavioural activity assisting in the construction of new ideas and engagement.

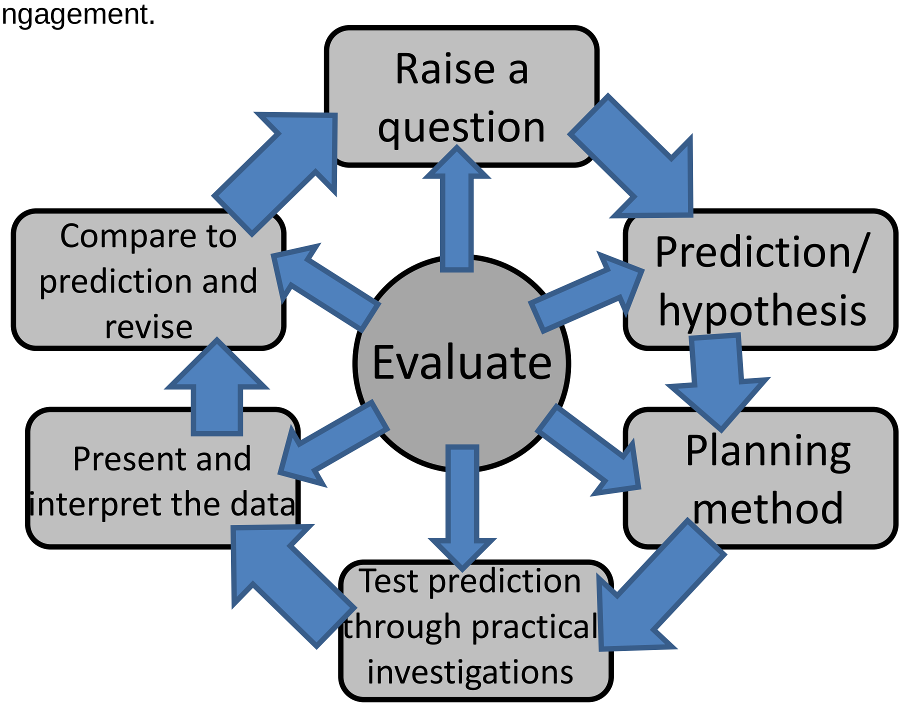


Fig. 1.2 Science investigation inquiry cycle

Activities will be undertaken in an environment of collaboration with peers with the teacher acting as a facilitator or expert. In these roles the teacher will be scaffolding the activities at a class, group or individual level. Using their expertise in teaching science and experience of pupils' ability to understand scientific concepts, they will aim to place activities (both knowledge and skills) in the pupils' 'zone of proximal development.' With skilled activity design and well-timed interventions teachers can promote the construction of conceptual understanding without allowing a passive transmission of knowledge. This will provide pupils with the opportunity to remain in control of their learning and so foster group and individual responsibility. Please refer to *Table 3.1* (See Appendix 3A) to see how this definition fits in with alternative models of teaching.

Table 1.1 Definitions of key terms for current research

Key Terms	**Summarised definitions**
Inquiry-based teaching	Opportunities provided for investigation of an authentic question about the physical world using practical work or secondary sources. Pupils are supported to construct knowledge and understanding in a collaborative environment with the teacher acting as a facilitator.
Authentic question	A scientific question linked to the real world which the pupils do not already know the answer to.
Practical Work	Pupils interact with materials or physical models that illustrate the scientific phenomena being studied
Inquiry/Enquiry	A pedagogical process framed as an independent, supported investigation.
Investigative science UK Scientific inquiry UK Inquiry cycle US	See *Fig 1.2* Question/hypothesis – predict- plan-experiment- present data- interpret – conclude by comparing to prediction
How Science Works/ Working Scientifically	UK National Curriculum – Scientific inquiry (as above) and the understanding of how scientific ideas develop, are evaluated and impact society.
Critical Thinking Skills	A reflective thought process where decisions are made and evaluated. Reasoning can be provided for decisions taken
Attainment	A measure for the level of scientific knowledge and understanding that a pupil can demonstrate independently

‘Ability’, as used in this dissertation, is common parlance in the teaching profession and although it is not well-defined from a researcher’s perspective, as a practitioner its meaning seems widely understood. At its simplest, in the setting for this research and at the time of writing, ability is a crude measure of a pupils’ academic performance based on two tests, their Key Stage 2 (age 11) English and Maths Standard Attainment Tests (SATs). The average level achieved by a pupil in these two tests is used across all subjects in the secondary school to place the pupil informally and qualitatively into a lower, mid or higher ‘ability’ band and quantitatively to provide them with a target level for the end of Key Stage 3 (age 14) and target GCSE grade for the end of Key Stage 4 (age 16). Once within school, and as some pupils do not arrive at the school with this data, these qualitative labels may change as the pupils develop but the quantitative target grade attached to them will remain until the end of Key Stage 4, as this is the measure that the government use to assess schools on pupil progress.

Clearly, from a researcher’s point of view this is a very crude measure as it does not take into account the complex set of factors that determine a pupil’s current or future aptitude for learning. In the context of this research, the ‘label’ is not based on any test of scientific knowledge or skills, nor does it take into account a pupil’s disposition to learn, or the social, cultural or educational factors that affect this. As the ‘ability’ measure of a pupil is based on the English and Maths SATs scores then it can be assumed that pupils in the lower ability band are likely to have weaker literacy and maths skills which may have an impact on their learning in science irrespective of any aptitude for science. Conversely, some pupils may perform well in these two subject areas, or be coached differentially to do so, and hence a high aptitude for science does not necessarily follow.

However, taking these points into consideration the concept of ‘ability’ as used in the dissertation is still a useful short-hand by which to categorise and therefore refer to pupils. It is the measure used by schools and the government in assessing pupil progress and therefore the performance of a school, and as such its use in this research piece has a high level of ecological validity. Weak or ‘low ability’ pupils are defined as those who enter the school with an average Key Stage 2 SATs score of Level 3 (L3) or below; mid- ability a L4; strong or high ability, L5. Where the pupils are taught in mixed ability groups for this research, this is defined as a class which consists of pupils from all ability bands although not necessarily in equal numbers. Due to the school intake the classes have a small number of lower ability pupils (i.e. between 0-4 children) and roughly equal numbers of mid and high ability pupils. However, as already discussed this does not mean that they always show attainment in line with these classifications.

1.5 Experimental Design and Research Methods

A mixed methods approach was used to investigate the three research questions and two sub-questions detailed in paragraph 1.1. This had the benefit of providing both qualitative and quantitative data for analysis. The first two questions produced a between groups intervention study, using a quasi-experimental design with pre- and post-testing of control and treatment groups. The intervention was a three topic scheme of work covering 24 lessons employing solely inquiry-based teaching strategies. The scheme of work was produced using methods and approaches listed in *Table 3.1* (See Appendix 3A).

The pre- and post-attainment tests used were the standard tests used by the research school science department for testing pupils' progress every three-topic cycle. They are comprised of 50% questions from the three topics most recently studied and 50% all previous topics. The pre and post-test were therefore different in content but equal in composition. Pupils' pre- and post-test scores were firstly compared individually to find the change, and then the change in attainment was compared between the control and the treatment group.

The pre- and post-critical thinking skills test used was Cornell Critical Thinking Test – Series X[1]. This was chosen as it is a peer-reviewed test that has been developed for pupils of the age included in the research cohort. The pre- and post-test were the same as problems associated with pupils carrying out the same test are smaller than the variables generated by producing two different tests. As with the attainment tests, pupils' pre- and post-test scores were compared and then the change in score was compared between the control and the treatment group.

Research question 3 was investigated using three focus groups, two pupil, and one teacher. The pupil groups comprised of volunteers from the control and the treatment group. The teacher group was made up of all the teachers that taught the inquiry-based scheme of work. The group discussion was carried out using a semi-structured interview format which produced qualitative data, that was analysed using the grounded-theory approach.

[1] http://www.criticalthinking.com/

1.6 Summary of Research Chapters

In Chapter 2 the literature relating to the study will be reviewed with the aim of clarifying the existing research, identifying gaps and exploring a range of perspectives within scholarly texts. The ideas will be distilled with the view to reaching a clear definition of concepts that can then be applied to this research.

Chapter 3 sets out the design of the study; an explanation will be given for the key aspects including the operationalisation of variables, analysis and the analytical approach that will be used for each of the three research questions. These choices will be justified in relation to alternative methods that could have been used ,with reference to literature concerned with similar variables to this research.

The results for research question 1 and 2 will be presented in Chapters 4 and 5. The data analysis will be presented as a comparison between pre- and post-tests scores and then the difference in any change found between control and treatment groups. In all cases this data will be analysed using appropriate statistical tests in order to ascertain the statistical significance of the findings. These findings will be presented following the statistical analysis and then where they will then be discussed with reference to literature alongside an evaluation of the methods used. Chapter 4 also includes a comparison of the teaching methods used in the control and treatment groups to generate data on the fidelity in the treatment group and to prove the intervention provided a difference to the standard teaching pupils received. A further analysis of the correlation between attainment scores and critical thinking skills scores is presented in Chapter 4 to question the interdependent nature of these two variables.

Chapter 6 will present the qualitative results gathered in answer to research question 3. This will include a thematic analysis of all three of the focus groups, which will then be displayed in a semi-quantitative format, with the number of comments per theme and then per individual tallied. Comparisons are drawn between responses from the control and treatment pupil groups and between the treatment pupil group and the teacher group.

The conclusions from all the research chapters will be presented in Chapter 7 where findings will be summarised and placed in the context of current literature. Major themes investigated will be commented on alongside suggestions for further research and advice for practitioners. The limitations of the study will also be discussed in terms of scope, methods and confounding variables.

The Appendices contain information on inquiry-based strategies and critical thinking skills to support the experimental design in Chapter 3. The Appendices associated with Chapters, 4 and 5 contain data in the mid-process of statistical analysis, referenced in the chapter, for the reader to refer to alongside that which is presented in tabular and graphical format within the chapter itself. There is also focus group transcripts and initial thematic analysis, which are summarised and displayed in Chapter 6, and therefore remain in the appendices to allow the reader to follow the process of analysis from the raw data.

1.7 Exclusions

Inquiry-based teaching is not specific to science but my study focussed on this as the subject matter lends itself to this style of teaching. Similarly critical thinking skills are not science-specific and this was highlighted by the use of a non-subject specific critical thinking test. This was a short-term and small study with no intention to produce findings that would be generalisable to other contexts.

1.8 Conclusion

To conclude this chapter, this is a study borne out of a practitioner's view that inquiry-based teaching could produce positive gains in pupils' attainment and critical thinking skills. The aim was to test this hypothesis using a mixed methods approach to gain the benefits of qualitative and quantitative data for analysis. In the next chapter will follow a more detailed analysis of the literature surrounding this subject which develops a discussion around the practitioner's perspective leading onto a considered experimental design.

Chapter 2 – Literature review

2.1 The Practitioners' Perspective

To develop my point from Chapter 1, the starting point for this research was my belief, that the use of inquiry-based teaching improves understanding, develops better thinking and motivates pupils. This strategy must have at its heart a philosophy for the most appropriate way to educate children; in this case constructivism seems a useful theory of knowledge . The status of constructivism referred to as a "secular religion" by Phillips (1995, p. 5) in the UK education system is certainly high ranking as can be seen from the tone of teacher training and professional textbooks. This is supported by studies into science teachers' beliefs (Wallace & Kang, 2004; Wilson & Mant, 2011; Brown and Melear, 2006; Minner, Levy, & Century, 2010) which conclude that teachers believe that involving pupils in the active construction of their understanding, through inquiry-based teaching strategies, results in a deepened understanding and development of skills other than scientific recall and therefore, in this research these ideas will be tested.

2.2 Competing Definitions and Importance of Inquiry-Based Learning in Literature

Although a definition of inquiry-based teaching has been reached for this research and is outlined in *Table 1.1.*A A discussion follows of the literature used to inform the definition of inquiry-based teaching given in 1.4.

There are differences in the definitions of inquiry-based learning between research articles, largely conducted in the US, and those in from the UK. Authenticity runs through the former, as well as autonomy (with scaffolding) and pupil devised questions (Crawford, Krajcik, & Marx,1999; Crawford, 2000; Hmelo-Silver, Duncan, & Chinn, 2007; Brown & Melear, 2006; Kuhn & Pease, 2008); whereas the British interpretation favours 'investigative work' and is more tentative in its regard to an open-ended structure. NESTA (2005) suggest that inquiry-based learning requires that only one part of the inquiry or investigation cycle (*Fig 1.2*) may be used at any one time. It also asserts that in the classroom it is unlikely that any investigation would fulfil the 'authentic' criteria of neither the student nor the teacher knowing the outcome; more likely is that students will construct new knowledge for them based on an inquiry-process into accepted scientific concepts. The NESTA definition also suggests that pupils may investigate phenomena with a prior knowledge of the result to be found;

this is in complete opposition to US research,mentioned above, where authenticity is fundamental.

Millar (2004) asserts that pupils may have some freedom in the design of investigative work but accepts that this cannot be considered ‘authentic’ as pupils are expected to construct schemas that are aligned with the agreed body of scientific knowledge. This viewpoint can be challenged only if authenticity is seen from the perspective of a pupil, who is meeting an idea for the first time. He also suggests that genuine open-ended, investigative work is valuable mainly to pupils who continue to study science at a higher level as the general public do not carry out the process.

White and Frederiksen (1998) concluded that questioning, predicting, experimenting, modelling and applying (Fig. 1, p. 5)can play a significant role in learning physics. . Hmelo-Silver *et al.* (2007) believe the skills pupils use when engaging in investigations are “collaborative learning and activity... [they are] cognitively engaged in sense making, developing evidence-based explanations, and communicating their ideas” (p.100). These examples show that the skills of scientific inquiry do not need to be transferred to a future career as they are used as tools to aid cognitive development in general.

Development of inquiry skills is an aim of the science curriculum that Millar and Osborne put forward in ‘Beyond 2000’ (1998),which seems to contrast Millar’s (2004) assertions Therefore, perhaps Millar’s point is that the current nature of investigative work in schools is not promoting the development of these inquiry skills, but rather trying to make unhelpful comparisons with industry as he suggests that “the argument that an understanding of the methods of scientific inquiry is practically useful in everyday contexts has been over-emphasised. For most purposes a systematic, common-sense approach will suffice.” (Osborne & Millar, 1998, p.11). Therefore the development of age-appropriate investigative skills becomes a curriculum design issue, where the skills that should be developed through the inquiry process should be emphasised over the direct links to science processes in the professional scientific community.

Despite the doubt cast on the relevance that teaching investigative science has for developing ‘scientific literacy’ (Millar & Osborne, 1998, p.13) it remains a prominent element in science education. Although inquiry-based teaching does not need to use the inquiry-cycle to be defined as such, it is true that for science the inquiry-cycle lends itself to the strategies that are listed in *Table 1.2.* Within this, UK research is dominated by improving the use of practical work is (Nuffield, 2012; ASE, 2009; NESTA, 2005) rather than the ‘inquiry cycle which is more closely aligned with inquiry-based teaching;as practical work can be embedded in the inquiry cycle but it is not a necessity. In the 2005 UK National Curriculum the term ‘How Science Works’ was introduced to broaden the impact of the ‘scientific enquiry’ strand, this was created “to reflect that ‘scientific enquiry’ was not just a set of experimental skills” (SCORE, 2011, p. 1). Abraham and Millar (2008) found that in the majority practical work was used to illustrate phenomena whilst development of skills required for the inquiry cycle were expected to be absorbed passively. Thus, without teaching emphasising the inquiry-based learning skills, Millar’s (2004) conclusion on the value of investigative work in practice is born out.

However, the argument for scientific inquiry being used to promote useful skills rather than model scientists (Millar, 2004) is in opposition to the National Curriculum (2013) “Working Scientifically” strand which details high level scientific skills to be taught (p. 4). Lunetta, Hofstein, and Clough (2007) question this curriculum requirement in that by expecting pupils “will come to the same understanding as scientists reflects a naïve empiricist view of scientific knowledge.” (p. 407). Both Lunetta *et al*. and Millar (2004) encourage inquiry-based teaching strategies but they warn about the use of practical work as an inductive method to generate understanding, as would be the case when using practical work within an inquiry-based pedagogy This danger is also highlighted by Wellington (1998), however in the same book Hodson suggests that a way to avoid this is to accept that pupils may not reach the same conclusions as scientists and as such pupils should be able to decide how they will carry out an investigation and to support this “teachers should seek to create a dialogic context in which meaning is co-constructed” (p. 101). In practice, ideally, pupils’ ideas are discussed with the teacher, questions are skilfully constructed and targeted information is provided when a pupil needs to reach the next stage in their understanding, in line with the strategies of inquiry-based teaching.

In continuing the discussion of inquiry-based teaching definitions, The Royal Society (cited in Science Select Committee, 2006) promote the use of open-ended investigative work but do not use the term ‘inquiry’ which brings ambiguity. The SCORE report (2011) compares the UK interpretation of inquiry skills with Alberta, Canada and Hong Kong finding that “the UK has a broader, more appropriate and modern interpretation of HSW” (p. 3). The Nuffield Foundation (cited in Science Select Committee, 2006) uses the terms “methods of science” and “investigation” (p.185) rather than ‘inquiry’. This variety of definitions as well as the priorities in science education for improved ‘practical work’ blurs the meaning of inquiry-based learning in the UK which creates difficulty in interpreting it for classroom use. Therefore, with little guidance, teachers must develop pupils’ inquiry within the context of the knowledge strands of the curriculum, and in many cases this has been ineffective (Lunetta *et al.* 2007; Millar and Abrahams, 2008).

The research outlined thus far demonstrates the complexities of defining this teaching strategy. Therefore, the literature used to generate the definition for this research *(Table 2.1)*), offers the basic principle that in order for 'Working Scientifically' to be taught through an inquiry-based teaching approach the curriculum needs to be well researched, explicit in it aims, scaffolded, with an emphasis on science literacy and context in order to produce a coherent set of resources that are transferrable to the classroom.

2.3 Link Between Inquiry-Based Learning Research, Practice and Pupil Progress

As summarised in the previous chapter (1.2) Hattie's (2008) Visible Learning presents results indicating that inquiry-based learning did not provide improvements in the 'zone of desired effects' (*see Fig. 1.1*). Hattie's definition of 'zone of desired effects' is based on pupils making better progress by a particular intervention than the average progress across all well-implemented interventions or changes. It also reflects the expected attainment progression for pupils in their school career.

Hattie's synthesis of four inquiry-based learning meta-analyses encompassing 205 studies gave an effect size of 0.31, (*See Fig. 1.1*),comparable to the average effect that a teacher would have on a pupil's attainment in a year. However, the most recent meta-analysis used was from 1996; teaching practices, interpretations of inquiry-based learning and professional training have progressed in the intervening years. Hattie counters this argument (2012) when he asserts that he accommodated further studies leading up to 2012 and the rankings of influence did not change,

However, in his discussion he does go on to report that effect sizes on scientific processes (assumed to be skills such as planning, data analysis and evaluation) were above 0.4 in two studies. In one, inquiry-based learning produced a 1.02 effect size on critical thinking skills, leading to the conclusion that '...inquiry-based instruction was shown to produce transferrable critical thinking skills' (Hattie, 2008, p. 210).This impact on critical thinking skills will be discussed and defined later.

2.4 Inquiry-based Teaching and Education Policy

Despite Hattie (2008) empirically devaluing the impact of inquiry-based instruction on pupil progress, the volume of research in this area indicates positive interest alongside its high regard from teachers (Wilson & Mant, 2011; Wallace & Kang, 2004). This is supported by the UK National Curriculum "Working scientifically" (Department for

Education, 2013, p. 4) and the National Science Education Standards in the USA (National Academy of Sciences, 1996). The UK National Curriculum definition mirrors that advised to state education boards in the US, where the inquiry process is defined as: development of a pupil question followed by collecting, recording and processing data ending with evaluation of results and identifying further work.

Both of these documents clearly state what should be taught; however within them there is no indication of how to meet these requirements. This creates a policy and implementation divide where the body of knowledge and skills to be taught is explicit, whereas the strategies to be employed by teachers or the processes used by pupils to acquire it are not. (Horner, 2011; Keys & Bryan, 2001).Therefore, it seems that the role of teachers in applying education policy is under-addressed considering the autonomy of the teacher in the classroom and hence their impact on the learning (Brown & Melear, 2006). Therefore, when designing the intervention for this research, the requirements of the "Working Scientifically" strand must be met, but the teaching strategies employed will be based on *Table 1.2.*

2.5 Teaching Inquiry or Through Inquiry: The Teaching Process

The difficulty in defining inquiry-based teaching arises in the separation of teaching the inquiry process as is expected by the National Curriculum and teaching using inquiry-based strategies. The expectation for teaching science in the UK is that the process strand will be integrated with the subject knowledge delivery. This narrower use of the term 'inquiry' has evolved by application to only scientific settings rather than other contexts. This adds confusion when promoting the use of inquiry-based teaching strategies which are far broader than 'science inquiry', and more in line with the general definition of inquiry. Consequently, inquiry-based teaching can be applied to all subjects as shown in *Table 3.1* (See Appendix 3A). In essence the science version of inquiry or the inquiry cycle *Fig 1.2* requires pupils to learn to think like scientists, In practice, this is only promoted when pupils are carrying out practical investigations (Ryder, 2011; Millar, 2004). These investigations are often asides to the main feature in UK teaching, that of developing knowledge.. the in these investigations is largely on reproducing expected results to support conceptual teaching rather than the inquiry process itself (Abrahams & Millar, 2008). In order to think like scientists, pupils need to be taught and provided with opportunities to develop, practise and refine the skills of inquiry in its broadest sense at frequent and regular intervals (Yeomans, 2011).

The ‘Working scientifically’ strand (DfE, 2013) states that pupils must be able to; “Ask questions and develop a line of enquiry based on observations of the real world…”(p. 4), therefore, pupils must initially be taught the scientific inquiry process.. It cannot be a straight-forward reproduction of real-world scientific practices as pupils do not have the cognitive abilities or experience to imitate scientists (Keys & Bryan, 2001). Yeomans (2011) and Ryder (2011) suggest that teaching the processes of scientific inquiry does not necessarily need to be through inquiry, this paradox is also highlighted by Kirschner, Sweller, and Clark (2006). They argue that expecting the same process from pupils as from experts is unrealistic and increases cognitive load to a point where it can lead to “misconceptions, inefficiency and frustration” (p. 79); this view is echoed by Kuhn, Keselman, and Kaplan (2000) as they attest that the use of under-developed skills could be “counterproductive” (p. 496). Tsai and Huang (2002) agree that pupils cannot be compared to scientists where “experts have well-developed or more integrated knowledge structures to help them solve problems” (p. 163). This agreement does not produce a solid argument against developing these inquiry skills, only an argument against assuming their presence in pupils.

Kirschner *et al.* (2006) disagree with the idea of teaching inquiry skills from the position of long-term and working memory. They assert that without prior knowledge of inquiry skills the demand on working memory would be too great to promote any kind of learning when required to use them. Forming an argument outside of the constructivist framework Kirschner *et al.* disregard the accepted position that pupils already possess, from prior school-based learning and experience in the world (Dewey, 1899), knowledge and skills that can be further developed. In support of Kirschner *et al.*’s ideas, it is frequently concluded that pupils have insufficient experience in the use of the scientific skills that are required for inquiry-based learning to have complete independence (Kuhn *et al.* 2000; Kuhn & Pease, 2008; White & Frederiksen, 1998), and in line with Piaget’s work, it is pupils’ biological development that may limit their progress (Keys & Bryan 2000). It is suggested that teachers and curriculum designers over-estimate pupils’ ability to complete these processes and cognitive limitations should be better considered (Kuhn & Pease, 2008; Millar,2004).

However, in contrast to Kirschner *et al.* the aforementioned researchers do not suggest a failure in the inquiry-based teaching process but advocate the use of the same dual strategy highlighted by Yeomans (2011) and Ryder (2011) where pupils initially need to be taught about inquiry before they can use it. The tasks must be in stages and carried out procedurally, and frequently, before pupils can have an appreciation for the links that can be made between them and the development of scientific understanding. The skills needed for inquiry must be broken down into smaller, more basic, elements (Lunetta *et al.,* 2007) through the use of “scaffolding” (Wood, Bruner, and Ross,1976). This may include teaching them by direct instruction (Yeomans; Ryder), but as a starting point teachers must be aware that the skills place a cognitive demand on the pupils in the same way as the learning subject knowledge.

A further issue which Kirschner *et al.* highlight is that some studies do not distinguish discovery learning, problem-based, experiential and inquiry-based teaching strategies from one another. Kirschner *et al.* do this under the title of the ‘Minimal-Guidance’ approach which they conclude has failed. This lack of distinction prompted Hmelo-Silver *et al.* (2007) to dispute their arguments by evidencing the positive outcomes that have been found in applying a “scaffolding” approach to inquiry-based teaching, this strategy was derived from Vygotsky’s (1978) conclusions on placing tasks within a pupils “zone of proximal development.” (p. 85). In a survey of literature used in this chapter, inquiry-based learning is closely identified with a constructivist ideology (Crawford *et al.,*1999; Ryder, 2011; Keys & Bryan, 2001; Klahr & Nigam, 2004; Kuhn & Pease, 2008; Minner *et al*., 2010) and so this theory’s absence from the work of Kirschner *et al.* highlights a weakness in their argument. Researchers collectively promote the idea of teachers as facilitators in inquiry-based teaching, providing structure, questioning and well-timed direct instruction (Hmelo-Silver, 2007; Keys & Bryan 2001; Crawford, 2000; Kuhn & Pease, 2008; White & Frederiksen,1998; Hodson,1998) this is in opposition to the Kirschner *at al.’s* definition of inquiry-based learning where pupils receive minimal guidance. Therefore, in line with the majority of research, scaffolding will be incorporated into the inquiry-based schemes of work that will be written for this research.

Indeed two meta-analyses by Lazonder and Harmsen (2016) and Furtak, Seidel, Iverson & Briggs (2012) also cite Kirschner *et al*.’s (2006) arguments as being representative of one definition of inquiry-based learning, that of offering minimal guidance. In both of these analyses evidence is synthesised to establish the effect of different elements of inquiry-based learning on learning measures. Whilst both research the degree of guidance received, student age is researched in the former, whilst it is cognitive features of an activity in the latter.

In contrast to Kirschner *et al*.’s assertion that inquiry-based learning has minimal guidance, both of these analyses synthesised research that was self-termed inquiry-based learning and hence papers included had applied differing degrees of guidance. In the case of Furtak *et al.* (2012) the papers selected were defined to be on a continuum from totally teacher-led to totally student-led and for Lazonder & Harmsen (2016) the type of guidance in each research piece was categorised from the most minimal, being the breaking down of tasks into smaller more manageable ones ("process constraints" p. 8) to "heuristic prompts" (p. 8) i.e. telling pupils what to do and how to do it. Furtak *et al.* concluded that the greatest positive effect size was seen when comparing teacher-guided inquiry with traditional didactic teaching suggesting that inquiry-based learning had the greatest effect on "student learning of science" (p. 322) when led by a teacher, although undefined guidance e.g. computer also had large effect sizes. This conclusion was in agreement with Lazonder and Harmsen who found that "…larger effect sizes were associated with more specific types of guidance" (p.23). These ideas lend themselves to the notion the inquiry-based learning can include prompts and teacher guidance and still be defined as such which has implications for the design of the scheme of work in this research piece.

In addition to the general findings of Furtak *et al.* (2012) on guidance, they further define "student learning of science" (p.322) by the cognitive features of inquiry-based learning and divide these into four categories "conceptual structures" i.e. body of knowledge and reasoning; "epistemic frameworks" i.e. evaluating and interpreting scientific evidence; "social interactions" i.e. how scientific knowledge is communicated and "procedural" i.e. ability to create questions and design experiments (p. 305). As such, inquiry-based learning had the least effect on "conceptual structures" which is the category most aligned with the attainment measure in this research piece; whereas gains were found more often in the epistemic framework, being the application of skills - a measure which may reveal itself through critical thinking scores or the qualitative data collected from both pupils and teachers.

This idea that there are different cognitive domains to be measured was also discussed by McConney *et al.* (2014) when reasoning why pupils who stated that they received high levels of inquiry-based learning tended towards lower PISA (2006) "science literacy" (p. 27) scores (PISA uses 'science literacy' as measure of scientific knowledge, understanding and application that is de-coupled from the curricula of a country). They assert that although this dimension of learning seems to be negatively affected, other areas of learning are not measured by this instrument such as "depth of understanding and development of ideas that mimic scientists' deep understanding of specific questions or topics" (p. 28), which are more in line with the epistemic domain cited above. As such they conclude that inquiry-based learning should not be written off but that the elements that promote learning as well as pupil engagement should be better researched and identified.

2.6 Constructivist Pedagogies

In order to support the definition (section 1.4) and the construction of an inquiry-based intervention, discussion turns to the ideological framework within which it will be set. Constructivism is based on the idea that knowledge cannot be transferred wholesale; pupils must construct their knowledge from sensory perceptions and then integrate it into their current schema. The connection between inquiry-based learning and constructivist ideology, as suggested above, is frequent and the association is often made without discussion (Miri, David, & Uri, 2007).

The constructivist process for teaching begins with knowing the baseline of pupils' understanding; teachers must then create differentiated learning experiences.; They must skilfully explain, question and informally assess, adapt and feedback so that pupils can develop their own understanding and thus learn effectively (Hattie, 2012; Wellington & Ireson, 2008; Ofsted, 2013). Inquiry-based learning is one of a number of strategies that can be used to meet these criteria and in addition it can support the development of a large variety of science-specific (White & Frederiksen, 1998; Kuhn & Pease, 2008) and generic cognitive processes e.g. “grappling with data” (Crawford, 2000, p. 926), “develop productive collaborative relationships with team members” (Crawford *et al.* 1999, p. 720), and therefore these constructivist principles will be used to assist in the development of the intervention for this research.

However, there is opposition to the constructivist pegagodgy, Phillips (1995) in his critique, acknowledges the ‘good' side of constructivism; its focus on pupil activity and social co-operation, but questions the wholesale acceptance of this ideology.. His observation that the ‘ugly' side of constructivism is its status as a “secular religion” (p. 5), can be supported by modern teaching resources e.g. camtools.cam.ac.uk; www.teacherstoolbox.co.uk; www.theschoolinthecloud.org; www.upd8.org.uk/wikid. This has led to the dampening of other teaching strategies which have proved to be useful tools e.g. Direct instruction (Hattie, 2008), passive learning through lecturing and memorising (Liu & Matthews, 2005) and rote learning (Willingham, 2009; Christodoulou, 2014). Phillips opposition is supported by Ofsted (2014) who state that there is no requirement to teach in a particular style or in line with a consistent methodology. Trends and fashions in teaching do exist, and Phillips' point does highlight the dangers of being blinkered to other ideologies and strategies. This is particularly important at present where political rhetoric has promoted a return to a more instructional content-focussed style (Burns, 2012; Walker, 2012). However, constructivism is a broad ideology which enables the development of understanding where many teaching strategies, including direct instruction, can have a place.

2.7 Inquiry-Based Teaching and Social Constructivism

Vygotsky's (1978) conclusions on the construction of meaning through language, are a key feature of learning in UK education where group work and discussion activities are frequently used (www.upd8.org.uk/wikid; Ginnis, 2001). He stated that "Learning awakens a variety of internal development processes that are able to operate only when the child is interacting with people in his environment and in cooperation with his peers" (p. 86). Although inquiry is a process which can be developed and used independently many inquiry-based learning activities in current use rely on social interaction to promote discussion. (www.upd8.org.uk/wikid; Addey, 1999; http://www.nuffieldfoundation.org).

However any assumption that all social interactions will benefit the construction of understanding is flawed. Wellington (1998) discusses the draw backs of group work in schools where pupils do not have the social skills to aid cooperative learning. Ofsted (2013) acknowledges some excellent use of group or pair work in practical tasks but recommends that opportunities for pupils to work on their own be incorporated.. Although individual work may offer benefits in areas such as autonomy, resilience and motivation, barriers to this such as resources, time and class size, often make group practical and research work preferable and consequently group work is often the default, rather than considered, tool used.

Crawford *et al.* (1999) concluded that components of inquiry-based teaching such as developing authentic questions, student ownership of an investigation and teacher facilitation produced more collaborative behaviour. They also discovered that it takes "up to 8 weeks to become fully productive in working together toward a group goal" (p. 720). This implies that when using group work, its positive effect is not immediate, as indicated by Wellington (1998), and if group work is to be a worthwhile experience then teachers must research methods, plan groups and monitor throughout the activity. Pupils of differing stages of social development must be assisted in interacting with peers effectively to ensure learning takes place.

This idea of thoughtfully planned group work is important for this research as there is a consensus in literature on the value of discussion and dialogic teaching (Mant, Wilson, & Coates, 2007; Lunetta *et al.*, 2007; Wellington, 1998; Hodson, 1998). It is also true that both students and teachers value group work (SCORE, 2008; Wilson & Mant, 2011; Murray & Reiss, 2005). In Murray and Reiss's research 48% of students responded that "having a discussion/debate in class" was a "useful and effective" way of learning (p. 4) this was the highest score recorded for a strategy in this question. In research garnering the beliefs of 'exemplary teachers,' discussion was highlighted as an important by teachers for developing higher order thinking as well as scientific understanding, this perspective was reflected in pupils ideas of exemplary teaching (Wilson & Mant, 2011); these surveys place an importance on social interaction that is in line with Vygotsky (1978). As such, incorporating discussion into the intervention for this research needs to be considered carefully and scaffolded appropriately for the age of the cohort.

2.8 Inquiry-Based Teaching and Practical Work

The place of practical work when teaching secondary science within an inquiry-based framework appears implicit, however it is not essential. In the varied definitions of inquiry, practical work is always at the core (White & Frederiksen, 1998; Minner *et al*., 2010; Kuhn & Pease 2008; Brown & Melear, 2006; Crawford, 2000) as it provides opportunities to test predictions, observe phenomena, manipulate equipment and collect data. However, challenges arise when ‘practical work’ is less clearly defined in practice. This was shown by the SCORE (2008) survey on teachers perceptions where definitions ranged from “laboratory procedures and techniques” (86%) to “presentations” and “surveys,” receiving 7% and 3% respectively (p. 9). This variety of perspectives complicates the idea that ‘practical work’ must be embedded in the inquiry-based teaching strategy.

Lunetta *et al.* (2007) offer a useful definition of ‘practical work’ that has been used by other researchers (SCORE, 2008; Dillon, 2008), it states; “learning experiences in which students interact with materials or with secondary sources of data to observe and understand the natural world” (p. 394) Millar (2004) and NESTA (2005) mirror this, but do not include secondary data in their version, thereby producing a narrower definition. However, NESTA do state that inquiry cycle can take place without the practical element as long as the aim is to “investigate scientific phenomena” (p. 15). By including ‘secondary data’ Lunetta *et al.’s* definition implies that they are in agreement with NESTA, practical work (i.e. manipulating physical objects) is not required for teaching the inquiry-cycle. . However, at a fundamental level practical work must take place in either a real or theoretical space to provide the secondary data so that the later stages of the inquiry process can take place. In this research practical work will be referred to as ‘an activity which involves the physical manipulation or observation of materials and/or models’ and therefore practical work will be used only as part of the design for the inquiry-based intervention.

Therefore, as practical work will be incorporated as a strategy in the intervention scheme of work it must needs to be well-planned. Abrahams and Millar (2008) found that in order for practical work to be effective, in promoting conceptual understanding. teachers must better plan in order to develop connections between the “domain of observables” and “the domain of ideas” (p. 1949).They advise that it must not be assumed that the ideas demonstrated through practical work will emerge from the data or observations, or that pupils will draw the same conclusions from data as scientists.. This point is agreed upon by many researchers (Millar, 2004; Lunetta *et al*., 2007; Hodson, 1998) and as such a teachers' role as an instructor must be assumed at the correct point in inquiry-based teaching, which will be determined by the progress of pupils. The instruction must be carefully considered to ensure it provides scaffolding rather than answers so that the practical work, in this research, would fulfil the requirement of inquiry-based teaching.

2.9 Critical Thinking Skills Definition

Many researchers (Ku, 2009; Renauld & Murray, 2008; Halpern, 2001; DCSF, 2008) cite the definitions of Ennis (1962,1985,1987,1993), whose ideas have adapted over the years to become "Critical thinking is reasonable reflective thinking, focused on deciding what to believe or do" (Ennis, 1993, p. 180). Much of Ennis' work is based around the creation of assessment instruments for critical thinking skills. Therefore his definition alongside his instruments and guide to assessment design in this area are frequently cited.

Ennis (1993) and Moseley, Elliott, & Gregson (2005) associate critical thinking with the top three levels of Bloom's Taxonomy; analysis, synthesis and evaluation, or in more modern usage; creating, evaluating and analysing. These skills also fall under the term 'Higher Order Thinking Skills' (Miri *et al.,* 2007) who also included application. These terms are common in research definitions and as such offer consensus, however Ennis (1993) asserts, that the terms are too vague and open to interpretation, and are interdependent rather than hierarchical. This leads to a more specific list of elements included in a number of definitions (cited in Fasko, 2003, p. 8), reflective questioning, metacognition, reasoning, decision making, evaluating and problem solving. Interestingly the DCSF (2008) separates critical thinking and creativity, as do Miri *at al.* (2007) but place both critical thinking and creativity under the heading of higher order thinking. This inclusion of creativity seems to be the one area of disagreement in literature as Ennis (1993) felt his original (1962) suffered from "excluding creative aspects of critical thinking, such as conceiving of alternatives, formulating hypotheses and definitions, and developing plans for experiments" (p. 180). Therefore, Ennis views critical and creative skills as interdependent and therefore the 'concrete' critical thinking skills outlined by the DCSF may not be achieved in isolation from the creative.

2.10 Importance of Critical Thinking

Teaching of critical thinking skills in school is minimal, despite teachers believing that pupils' ability to think, discuss and reason is important (Wallace & Kang, 2004; Wilson & Mant, 2001). Ennis (1993) suggests that if critical thinking is not tested then teachers will not develop it, which would explain its under-representation in curriculum design. The curriculum has become narrowed to teaching for the test; and as critical thinking skills are not tested it follows that they are not a priority. Many researchers (Ku, 2009; Renaud & Murray, 2008; Miri *et al.* 2007; Pithers & Soden, 2000; Arter & Salmon, 1987) link critical thinking skills to functioning well in society, which requires the ability to think critically. Marin and Halpern (2011) agree with conclusions from the National Committee Report (1996) that "Above all the country must enable people...to equip themselves for a world of work which is characterised by change" (pp. 4.2). Employers are recruiting in industries that are evolving faster than a generations' career span and as such some of the most important skills are those, such as critical thinking, which are transferrable.

2.11 Critical Thinking and Inquiry-Based Teaching

The link between inquiry-based teaching and critical thinking skills can be seen when comparing *Table 3.1* to *Table 3.2* (See Appendix 3A and 3B), as many of the skills in *Table 3.2* can be compared to the teaching strategies in *Table 3.1*. Halpern (2001) supports this idea when she asserts that "the active involved nature of laboratories would be expected to foster gains in critical thinking" (p. 283), although she also states that there is little evidence to support this. Miri *et al.* (2007) found that inquiry-orientated experiments was one of three effective strategies which improved critical thinking scores, particularly in the skill areas of inference and evaluation. However, Pithers and Soden (2000) suggest that when teaching critical thinking, it should be integrated into all teaching methods, as the aim is not for pupils to produce perfect answers but to develop the thinking process behind them. This suggestion is supported by Marin and Halpern (2011), and under different titles, Halpern (2001) as "better thinking" (p. 283) and Arter and Salmon (1987) "good thinking" (p.1) and finally Ennis (1993) as "reflective thinking" (p. 180). Therefore, pupil critical thinking skills will be assessed in this research, before and after the inquiry-based intervention, to conclude on its impact in this area.

Research suggests that critical thinking skills take longer to develop than subject-specific ones. Renaud and Murray (2008) tested college students' critical thinking skills using a non-subject specific test before and after a short reading task. There was no difference in the gains between the experiment and control group. This conclusion is supported by Halpern (2001) when she states that "Cognitive growth is a gradual and cumulative process; there are no quick fixes" (p. 273) and Ennis (1993) as he asserts "Learning to think critically takes a long time. Much reflective practice with many examples in a variety of situations is required." (p. 181). Even though the research carried out by Renauld and Murray may be flawed by its intervention choice, the results suggest that a longer term intervention may be more effective at showing gains in critical thinking skills.

2.12 Assessment of Critical Thinking Skills

Halpern (2001) provides a confident assertion that *Table 3.2* (See Appendix 3B) detailing a collation of critical thinking skills is representative of other researchers definitions (p.272). Her research was based on college students and although many are accessible to lower secondary pupils (as in this research), some need a simplified definition, as such her ideas have been reproduced and applied to a lower secondary school level curriculum (*Table 3.2*).

Current assessment of critical thinking is largely confined to college students, with some tests being developed for pupils of KS4 age. Research into whether these assessments should be subject-specific or generic in order to provide the greatest validity is available. Although, Renauld and Murray (2008) showed no critical thinking gains using non-subject-specific test in their research, small gains were found when using a subject-specific test. However, they also suggest that a general measure test may be more useful over a longer intervention period and also, as some pupils will not have the knowledge base to be tested with a subject-specific test, another variable would not be introduced into the assessment.

There are two options for test formats within critical thinking; multiple-choice and essay. Both Ennis (2003) and Ku (2009) suggest that multiple-choice is less valid than essay as they do not capture the disposition of the taker which is reported to have an effect on the score (Ku, 2009; Renaud & Murray, 2008; Miri *et al.* 2007; Ennis 2003; Pithers & Soden, 2000). Ku also suggests that the multiple-choice format does not reflect the nature of critical thinking skills; this position is supported by Halpern (2003) who states that 'multiple-choice critical thinking tests do not directly act as indicators of test-takers ability to think critically in unprompted contexts' (p. 356). Ennis also asserts that the generic nature of the tests means that the test-writer and taker will have "differences in background beliefs and assumptions" (p. 181) and therefore, may justifiably come to different conclusions; he also suggests that multi-choice tests often lack comprehensiveness and so miss the assessment of important critical thinking skills.

Both Ku (2009) and Ennis agree that essay-based tests are more comprehensive and valid, allowing the assessment of disposition and responses given without prompts. Ennis's argument, therefore, for the value and continued use of multiple-choice tests, particularly in the US, is that they require limited marking expertise and time. It is also the case that they can be used by younger people, as a much lower level of literacy and motivation are required to complete them, thus making this format the most practical for a lower secondary school.

2.13 Conclusion

This concludes the review of literature on topics most pertinent to this research. Inquiry-based teaching is a widely researched topic with little consensus. Definitions differ as much as the contexts and practical strategies that have been applied to research, which makes agreed conclusions challenging. Further complication is added when applying this pedagogy to science where other factors must also be considered, such as curriculum requirements and practical work, another area where research and conclusions differ. Therefore, this review aimed to give the rationale for the choices of key term definitions set out in *Table 1.1,* which then go on to form the framework for the experimental-design detailed in the Methods chapter which follows.

Chapter 3 – Methods

3.1 Introduction

In this chapter there will follow a methodology overview in the context of the research questions. This will lead on to a description of the methods used in relation to all aspects of sampling, administration, instruments, timeframe and resources. Precedents will be provided for this methodology in similar research and then discussion will then turn to the justification and rationale of the research project methodology as a whole and for the data collection instruments. The chapter will conclude with a discussion of the potential generalisability of the findings and conclusions.

3.2 Methods overview

This was a quasi-experiment using treatment and control group from Year 8 pupils, age 12-13 years old from a single school. Their pre- and post-test scores were analysed for both attainment and critical thinking skills before and after teaching of a three-unit scheme of work (the intervention). The category of action research is also applied as the researcher also teaches one class and is looking at the impact of the intervention on pupil outcomes through a variety of data collection methods. The data generated was both qualitative and quantitative: the quantitative data from the instruments used for RQ1 and 2 were used to test hypotheses through statistical analysis to compare control and treatment groups. The qualitative data provided by RQ3 added depth to the discussion by discovering internal themes and a semi-quantitative analysis was used to evaluate these themes and triangulate them against the data for RQ1 and 2.

At this point a re-iteration of the research questions is useful with an explanation on how each will be addressed.

3.3 RQ1 -What is the impact of an inquiry-based scheme of work on pupils' attainment?

Pupil attainment is a measure of their scientific knowledge and understanding assessed by the National Curriculum level achieved in a department-generated test, therefore the change in their level from pre- to post-test will be the measure of attainment in this research. In the research school pupils are assessed three times a year, the tests are constructed using questions from past Key Stage Three SATs (Standard Assessment Tests)[2] via the computer programme Testbase[3]. They are designed in two tiers, a lower level 3-6 paper and a higher level 4-7 paper; pupils with

[2] © Qualifications and Curriculum Authority

[3] http://www.testbase.co.uk/sec/index.php

a higher target for lower secondary science (Level 6 and 7) sit the higher tier and those with a lower (Level 4 or 5) vice versa. In terms of the research question stated above, the test that all pupils completed at the end of the previous cycle of three topics was taken as their pre-test score and the one at the end of the intervention period constituted their post-test score. Thereby, the change in their levels was used to measure the effect that teaching using inquiry-based strategies had on attainment from comparison of the treatment and control group.

RQ2 -What is the impact of an inquiry-based scheme of work on pupils' critical thinking skills?

Pupil critical thinking skills are not systematically assessed in the UK or at the research school; therefore a suitable instrument was needed. The Cornell Critical Thinking Test-Series X (CCTT) was chosen and purchased from the Critical Thinking Skills Co.[4] as a computer programme which the pupils could complete using individual passwords. Pupils sat this test immediately before and then again immediately after the intervention was concluded, thus creating their pre-and post-test score. The change in the score was used to measure the effect that teaching using inquiry-based strategies had on critical thinking skills from comparison of the treatment and control group.

RQ3 -What are pupils' and teachers' perceptions on the use of inquiry-based teaching strategies in science?

The perceptions of pupils and teachers were gathered through voluntary (opt-in) focus groups. Pupils needed written consent from their parents to be involved. There were two pupil focus groups, a control and treatment. They took the form of a semi-structured interview where pupils were asked the same series of questions in both groups to ascertain differences in response and to generate themes. The teacher focus group used the same format but included only the teachers who taught pupils in the treatment group.

3.4 Description of the Method

School and participant profiles

The school used for this research is the employer of the researcher; it is a larger than average Roman Catholic urban 11-18 school. Most students are of White British heritage and fewer students than average are eligible for free school meals. The proportion of students with special educational needs and/or disabilities is below the national average. (Ofsted, 2013).

[4] http://www.criticalthinking.com/

The cohort used for this research was 12-13 year olds (Year 8). The research period ran from February 2015 to June 2015, including data collection from focus groups. The cohort was taught science in six mixed-ability form groups for three periods of one hour, per week.

The teachers were chosen for two reasons; firstly, they were willing to teach the scheme of work exactly as prescribed and secondly, they expected to be present for the entire research period.

Pupils were assigned to control and treatment groups depending on the teachers that had been allocated to them at the beginning of the year. Pupils were not aware that they were part of a research project in terms of the teaching being received.

Other contributors to the research project were as follows: the Head teacher who had given consent both informally before the start of the project and formally in regards to ethical consent; the science technicians who provided the resources and equipment as detailed in the inquiry-based scheme of work each lesson; the IT technicians for the CCTT software installation and generation of individual passwords for pupils.

3.5 Inquiry-Based Scheme of Work Design and Production

The scheme of work (See Appendix 3C) to be followed by the teachers of the treatment group was designed to teach the content outlined in the National Curriculum for microbes, earth science and magnetism using inquiry-based teaching strategies (See Appendix 3A). Teachers were provided with an activity guide and teacher notes along with resources and a technicians guide.

In my role as a researcher I found that there was not a published or reviewed inquiry-based scheme of work that could be used as the intervention for this study. Although there are resources available to practitioners they did not fit with the school requirement for content and timeframe, and neither did they fulfil the definition of inquiry-learning that was laid out for this research. Therefore, as the researcher is an experienced practitioner the scheme of work was written exclusively for this piece of research (See Appendix 3C). The level of guidance used fell into the first three categories defined by Lazonder and Harmsen (2016) depending on where a pupils' 'zone of proximal development' fell. It can be seen in Appendix 3C that teachers were explicitly instructed to give prompts through timely interventions; it is also apparent from the scheme of work that the intervention had "process constraints" (p. 8) as tasks were both compartmentalised by lessons and then divided into small sub-tasks. The result of this means that the learning would be termed as 'teacher-led' overall with 'student-led' activities within, as researched by Furtak *et al.* (2012). The scheme of work aimed to develop all of the cognitive features defined by Furtak *et al.* with the assumption that they would all contribute to either a development in pupils' scientific understanding, measured via the attainment test, or in their critical thinking skills.

3.6 Administration

Teaching strategies and measures

All Year 8 teachers were briefed before the start of the research to ensure that they understood its aims, timescale and requirements. They all administered the pre- and post-attainment and critical thinking tests in accordance with the procedure detailed below. They also taught the three topics over the same period of time (24 lessons). The teachers of the control group taught using their normal teaching methods.

Consent was requested from all teachers for two lessons to be filmed during the research period using IRIS cameras and software[5]; all but one teacher consented. Filming took place in the 3rd and 7th week.

Attainment Measures

The pre-test for attainment and critical thinking was completed in the week following February mid-term 2015. The attainment test was administered by the teachers within a single lesson in the classroom; it was a 54 mark test, completed in silence. Before marking the test the class teachers photocopied three copies and gave them to the researcher to use for moderation. The testing procedure was repeated at the end of the experimental period (24 lessons). Pupils were anonymised, from their original data and scores on departmental spreadsheets, before the data was analysed using SPSS software.

Critical Thinking Measures

The critical thinking test was atypical for the school, accordingly all parents were sent an opt-out form, none of these were returned and so the entire cohort was eligible to take the test. The CCTT was originally set to be completed on the computer through software provided by Critical Thinking Skills Co. however when this was administered for whole classes the software crashed. Therefore, the company sent a PDF copy of the test which was photocopied for pupils and an answer grid created. Pupils were advised that the test was to find out about how they thought and that it was not about science. The test manual states that it need not be in a silent or formal setting and so instructions were given to teachers that the first section should be read out with the two examples at the start completed together as a class. Individual reading support could be offered to pupils with low literacy levels by teachers and teaching assistants

[5] www.irisconnect.co.uk

by reading out the question and possible answers. Pupils were also orally encouraged to complete the test when they seemed to be losing motivation.

The test was undertaken within a lesson and any that were incomplete were removed from the data analysis. This procedure was repeated at the end of the experimental period to provide a post-test score. The CCTT manual provided outlined the formula to be used to calculate the pupils' critical thinking scores (CCTT Score = c – ½i where c = correct and i = incorrect). Only pupils who completed both the pre-and post-test had their scores transferred to a spreadsheet anonymised and then analysed using SPSS software.

3.7 Pupil and Teacher Perception Measures

All pupils were invited to take part in the focus groups by the researcher. They were asked within their science lessons if they would like to give their thoughts on teaching in science and how they learn. Those interested were given opt-in consent forms to be signed by parents. 62 forms were taken and nine were returned, six from pupils in the treatment group and three from the control group, all of these pupils were included in the focus groups. In the focus group pupils were asked questions in a semi-structured interview style where some prompts were given and pupils could converse both with the researcher and each other; every pupil was given the opportunity to answer every question and were prompted to do so (this was also true of the teachers). They were informed that they would be anonymous in the research and that their answers would be recorded so that a written transcript could be produced (See Appendix 6A and 6B).

All teachers bar one (who had left the school) gave written consent and took part in the teacher focus group (See Appendix 6D). The grounded-theory methodology was applied to the qualitative analysis, the transcripts were colour-coded by the themes that initially rose from the treatment group and were then applied to the control group and then both were revised in light of the other. Each mention of a theme by a pupil was counted, to produce a semi-quantitative comparison between the control and treatment pupil groups; supporting quotes were also extracted (See Appendix 6C). The frequency of each thematic mention was calculated by dividing the total number of mentions per theme by the number of pupils in the respective focus group, this allowed data comparison between the control and treatment groups as they were different sizes.

The teacher transcript was colour coded with its own themes (Appendix 6D); these were tallied and a table was produced with an overall count and supportive quotes. This total count was then used to calculate frequency as described above. The teachers' themes were different to the pupils but similar pupil themes were put alongside teacher themes (Appendix 6E) so a comparative analysis could be carried out.

3.8 Precedents for Methods

There are a number of examples from literature that outline a similar use of the methodology and instruments used in this research to determine the effect of either inquiry-based teaching or similar interventions on both attainment and critical thinking.

The measure of attainment in this research is matched by that conducted by Mant, Wilson, & Coates (2007) where they "developed science lessons that had more practical work, more discussion, more thinking…" (p. 1707). The impact of this was found by measuring the proportion of pupils achieving level 5 in Key Stage 2 SATs tests before and after the intervention in control and treatment schools. This current research continued to mirror their methods by generating qualitative data from pupil focus groups, although theirs were only held in the treatment schools. In another similar piece of research by Geier, Blumenfeld, Marx, Krajcik, Fishman, Soloway, & Clay-Chambers (2008) state-wide standardized tests were used to measure the effect of an inquiry-based science curriculum on 7th/8th grade student outcomes. This is the same age cohort and similarly their pre- and post-test scores were compared to the control group.

In terms of critical thinking measures both Ernst and Monroe (2007) and Daud and Husin (2004) used The Cornell Critical Thinking Test Series X to measure the impact on critical thinking skills of an environment-based education and computer-aided extended reading classes respectively. In the former the age of the cohort was 14-15 years old (9th grade) and the latter the students were intermediate learners of English. In both cases the test was applied within a quasi-experimental design and was used in its entirety as both a pre- and a post-test.

The use of pupil focus groups alongside quantitative data from these tests was adopted by Mant *et al.* (2007), where pupil groups were involved in a semi-structured interview format; similarly this method of data collection was used by Miri, David and Uri (2007) when measuring the impact of 'purposefully teaching for the promotion of higher order thinking skills' (p. 353) on pupils' critical thinking skills. Miri *et al.* also made use of recorded lesson observations to provide triangulation of findings.

3.9 Research Rationale and Justification

This was an experimental study as the independent variable of either receiving teaching using inquiry-based strategies or in the usual style of the class teacher was determined and applied by the researcher. As the control and treatment groups are made up of pre-existing teaching classes the study falls into a quasi-experimental design and is more practical than a true-experimental design as the disruption of pupils' groupings would have not been possible in this school context. This approach is used as the effect of the intervention was measured with both pre-and post-test scores in non-equivalent control and treatment groups for attainment and critical thinking; these produced quantitative data. This study design aimed to minimise the effects of the teachers' and researcher's beliefs and provided the opportunity to test a number of hypotheses through the use of statistical tests. The collection of qualitative data provided an opportunity to discover themes and group beliefs which may suggest explanations for the findings from the quantitative data, providing a deeper understanding of the conclusions drawn.

Research duration

The research duration was chosen to be long enough to show small changes in attainment, which was supported by previous data on average attainment changes for pupils between tests. The timespan also aligned with the departments' schemes of work, the scope of this research and was acceptable to the school and teaching staff involved. In the literature, duration of an intervention when investigating critical thinking and/or the impact of an inquiry-based scheme of work on attainment ranged from a single piece of work (Renaud & Murray, 2008) to one and two year studies (Miri *et al.*, 2007; Geier *et al*., 2008; Mant *et al.*, 2007), but a number reflected this research in lasting over one section of work (Ernst & Monroe, 2007; Marin & Halpern, 2011; Khlar & Nigam, 2004). As with the studies outlined, a longer period may be advantageous for further research as it would reveal larger changes in attainment in both the treatment and control groups and so minimise the effect of anomalies. It would also give more time for teachers and pupils to adapt and refine their teaching and learning using inquiry-based strategies. This can be seen by the work carried out over these longer periods (Miri *et al*., 2007; Geier *et al*., 2008; Mant *et al.*, 2007) who all found critical thinking or attainment gains, and where a shared feature was that teachers received training and were involved in the development of resources; training was not possible within the scope of this research.

Research Topics

The topics chosen to be taught using inquiry-based strategies were determined by the department's unit of work timeline; thus not by researcher intent or for ease of delivery in this style and so the impact of researcher bias was reduced.

Research Cohort

Lower school secondary science is delivered in Year 7 and 8 in this school, all other year groups are preparing for national tests and so changes to their usual learning experience would not be desirable, therefore the research cohort became Year 8. The time of the year was also determined by the department timeline; advantageously by February in Year 8 the class had been taught science (and most other lessons) together for a year and a half so they knew each other well enough to interact in teacher-created small groups, although the disadvantage of this was that other social factors may have emerged that affected interaction or effective learning. In addition they had been taught in secondary school for a year and a half and by their class teacher for at least six months and so qualitative data collected on changing teaching strategies and their perception of these was more informed and useful. The structured assessment programme in Year 8 also lent itself to a pre- and post-test experimental approach with little disruption. The mixed ability nature of the classes in Year 8 also created representative control and treatment groups to which statistical analysis could then be applied.

Attainment Measures

The test materials used in the school, as outlined above, are common to many schools and have precedents in research. Published resources and software to support pupil assessment in lower secondary science are designed to report according to the levels produced by the QCA[1]. The use of the school's own department-generated tests would be both valid in context, as this is the usual attainment measure, and offer reduced disruption to pupils usual learning experience, additionally it also produced a large sample size (153 pupils) as consent was not required. In terms of the attainment research question the use of these standard tests meant that findings and conclusions drawn from this data would be true to the context of the research i.e. attainment in lower secondary school science in the UK, and so provide an opportunity for recommendations on the use of inquiry-based learning to be provided and relatable to other schools. Furthermore, the researcher was not the author of these tests thereby removing an element of bias. Although it is true that all teachers are aware of common questions that arise, no teacher involved in this research could have taught with knowledge of these specific tests.

A drawback of these tests was their two-tiered nature, as in order to produce a comparable attainment score the raw score required levelling. A calculation is usually used to provide 3 sub-level divisions i.e. Level 3 = 3c (3-3.32), 3b (3.3-3.66) or 3a (3.67-3.99). Although this provides the precision required for school attainment measures it is not as fine as the raw score. The research duration suggested that any attainment changes expected would be small i.e. less than a sub-level. Therefore in this research the levels were divided by ten rather than three from the raw data, thereby small changes in attainment could be distinguished. Moreover, small differences in attainment change between the control and treatment groups became apparent and available for analysis, providing quantitative findings to support conclusions drawn.

As already discussed, the use of standardised tests is common in this research area; however, there are examples of researchers creating their own instruments to measure learning gains. The flow-map analysis (Anderson & Demetrius, 1993) used by Bischoff and Anderson (2001) and Wu and Tsai (2005) was applied to an age range of 10-15 years old in these two research pieces and was favoured as it provides evidence for a deeper subject understanding by analysing connections pupils have made between areas of knowledge recall. This was an instrument of interest due to its focus on scientific understanding rather than recall. However, the attainment tests chosen have validity in this research where the research question is testing 'attainment' and not deep understanding; they are also suitable due to their previous use, context and production of standard quantitative data, thereby providing relatability.

Critical Thinking Measures

The Cornell Critical Thinking Test Series X is designed for ages 10-18 and measures critical thinking in four categories; induction, deduction, credibility and identification of assumptions. There is precedent for its use in this type of research and age group as discussed above; in addition it has been extensively peer-reviewed (Hughes, 1992; Malcolm, 1992). Suggestions for caution in this evaluation were applied to the current research, such as not using section scores for comparison due to small numbers of questions in each category and avoidance of comparisons with the normative data as the context of the data sets are ill-described. As the work of Ennis (1993) informed the definition used in this research for critical thinking skills, Hughes asserts that for content validity "If one accepts Ennis's conceptualization [of critical thinking], one would probably agree that the items adequately sample these skills." (p. 242), and therefore this is the case for this research. There is also a possibility suggested that reading ability may affect a pupils' ability to answer the questions however this test does state appropriateness for ages 10-18 and the pupils in this research are 12-13 so in most cases the reading age would be appropriate.

In the report Hughes (1992) cautions against the test's use for individual pupil judgements but states "that it may be an appropriate choice for evaluating instruction designed to improve critical thinking skills…" (p. 243), which is in line with this quasi-experiment. With this approach in mind Ennis (2003) suggests that there would be more problems with using two different forms of a critical thinking test for a pre-and post-test measure in terms of comparability of construct validity than those that would arise from the pupil familiarity with questions in using the same test.

One of the reasons for this is that the test is non-subject-specific and therefore definitively correct answers are not apparent to the pupils and could not be revised or looked up after the pre-test. Another reason for the choice of a non-subject-specific test was because it did not rely on pupils' scientific knowledge, it was accessible to all and less intimidating to lower-ability pupils; this was chosen to encourage higher rates of pupil engagement and test completion. Despite Renaud and Murray (2008) finding that for undergraduates, shorter term gains (within 90 minutes) were more detectable with a subject-specific test, they also suggested that a general measure test may be more useful over a longer period after more exposure to skill-based situations. It is also suggested by Ku (2009) that the multiple-choice measurement format does not reflect how critical thinking skills are used, often interdependently and without prompts. This is a position with which Ennis (2003) sympathises and suggests that an essay-based test would be more comprehensive and valid but understandably time and resources are factors and additionally, with this research, the pupils' literacy levels, determine that the multi-choice format is the more practical.

Use of this instrument removed experimenter-bias from its design; it was constructed and reviewed by other academics. However, experimenter-bias may challenge the validity of scores achieved by the researcher's class; her knowledge of the critical thinking skills being measured may have affected her delivery of the intervention leading to a greater development of these skills in the pupils taught compared to the other treatment classes.

Pupil and Teacher Perception Measures

A semi-structured interview design was chosen over structured as it allowed the researcher to intervene and question a theme in more detail and for the respondents to elaborate. The advantage over the unstructured style is that it provided prompts which were needed to keep participants discussing themes relevant to this research. The grounded-theory approach was used, Oktay (2012) justifies its use as it sits between a broad narrative analysis, which is used for holistic aims, and pre-determined categories which were not suitable to this research as this "… increases the likelihood of "forcing" the data into pre-existing categories that are not grounded in the data" (p. 54). Triangulation was provided by comparison of frequency data of teacher and pupil perceptions and the coding of themes which allowed further comparisons with the quantitative data from RQ1 and 2.

3.10 Fidelity

As the control and treatment group consisted of three teaching classes each it was important that there was fidelity in a number of areas. Firstly it had to be shown that there was an actual difference in the teaching received by the control and treatment groups. The recorded lessons, as detailed above, were watched at the end of the research period at five minute intervals. Five emergent categories of activity were created to facilitate statistical analysis through the use of tally charts. This method was used to try to remove bias and produce an objective and unambiguous record. However, the researcher's knowledge of the treatment scheme of work may have affected how each point in time was viewed compared to the control group recordings where the context of that time reference was unknown. The five minute data points were chosen as, despite drawbacks in their representation of the entire lesson, they do provide an objective data collection tool, in real time, providing ten points per lesson.

The recording of lessons was also used alongside pupil focus group comments to demonstrate fidelity in the treatment group, by providing evidence that teaching used the inquiry-based strategies outlined in the schemes of work.

The photocopied attainment tests given to the researcher for each class were for moderation purposes. The three tests were marked and compared to the class-teacher-marked versions. Fidelity was expected as the mark schemes for these types of tests are very clear and leave little to interpretation; all of the moderated tests were within one raw mark of the class-teacher marked copies. This process was not required for the critical thinking test as all class data was input by the researcher.

3.11 Generalisation

It is not possible for the conclusions drawn from this research to be generalised to offer definite recommendations due to the range of context-specific variables present. The pupils are from a similar socio-economic background as outlined in the school profile and are typically higher achieving than the national average which differs from other schools and Year 8 cohorts. This profile also varies between year groups and so making recommendations even within the same school would need to be applied with caution.

It is also true that the duration of the research only offered small gains, findings from research carried out over a longer period of time could offer a greater generalisation and scope for recommendations. It is clear from other research into this area, detailed above, that training teachers in the purpose and administration of inquiry-based learning was key to teacher engagement and fidelity; this was an area that was not dealt with in this research which reduces the relatability of the conclusions to other research. Despite these drawbacks there are elements of the findings and methods that could be relatable to other researchers, teachers or schools.

The tests provided consistency between classes and produced data with limited bias and a clear correlation between pupil pre- and post-test scores suggesting high levels of validity. This implies that similar instruments could be used in different cohorts or schools for a quasi-experimental approach. The design of an inquiry-based scheme of work written by a teacher was possible within the confines of the national curriculum and the timings provided for lower secondary school teaching of science, therefore this element of the research could be reproduced in other schools which may draw similar conclusions.

Chapter 4 - Data analysis, findings and discussion

4.1 In this and the following two chapters I will present the data for my three research questions. This chapter will present the data for research question 1 and 1a as detailed below; it includes a findings, analysis and discussion section. The data analysis comprises of;

Section A - A comparison between the teaching received by the control and treatment groups
Section B - An analysis of correlations between pupil prior and target attainment against critical thinking scores;
Section C - A comparison of critical thinking scores comparing pre-test to post-test scores
Section D - A comparison of the change in critical thinking scores between the control and the treatment group.

4.2 Research Question 1: What is the impact of an inquiry-based scheme of work on pupils' critical thinking?

Research Sub-question 1a: Is there a delivery variable that influences these effects?

In analysing the following data the probability (P) value that all of the statistical findings will be measured against is 0.05. If $p \geq 0.05$ then there is an equal or greater than 5% probability that the findings are due to chance and therefore the null hypothesis will be accepted.

Section A - Comparison of teaching strategies used in the control and treatment groups

Within the school setting the practicality of designing a replica control group was a challenge. The pupils were taught the same topics in the same amount of time but the teachers of the treatment group were permitted to teach using any method that they would normally employ, which may include inquiry-based strategies. Therefore, this analysis of strategies used was applied to see if there was a clear difference in the teaching received by the control and the treatment group. A tally was recorded every five minutes and a tally was placed in one of five categories that the teaching strategy fell into (discussed in section 3.12.1).

4.3 Analysis

Null hypothesis: There is no significant association between the teaching strategies and their use in the control and the treatment group

In order to test for association a *chi-squared test*[6] is used; if the *p* value is lower than 0.05 then there is a significant association between the teaching styles employed and the group observed i.e. control or treatment groups.

Table 4.1 Shows the teaching style that was taking place at five minute intervals; the style fell into five categories.

Teaching strategy	No. of tallies % of occurrence out of total tallies within group	Control	Treatment
Teacher-led	Tally % of group total	35 62.5	15 13.8
Individual work	Tally % of group total	9.0 16.1	10 9.2
Group discussion	Tally % of group total	0.0 0.0	38 34.9
Practical work	Tally % of group total	11 19.6	25 22.9
Researching a question	Tally % of group total	1.0 1.8	21 19.3
Total	Tally % of group total	56 100	109 100

Table 4.2 Shows the chi-squared test and its significance

	Difference between control and treatment groups
Degrees of freedom	4
Chi-squared value	58.7
Critical value	9.48
Significant	Yes

4.4 Findings

There are some clear differences in the number of instances of teaching strategies used between the two groups (*Table 4.1*). There were 35 counts of teacher talk in the control group which accounted for 62.5% of the points recorded for that group compared to 13.8% in the treatment group. This shows a considerable difference in the amount of time that pupils spent listening to their teacher or being involved in teaching that was being led from the front of the classroom. This is in opposition to the number of instances of pupils being observed in a group work where 0 instances of

[6] Ordinal data, in the form of a tally of teaching strategies used, and testing two independent groups.

group work were recorded in the control group compared to 35 instances in the treatment group, which accounted for 34.9% of all points recorded for that group.

A clear finding is that pupils in the treatment group spent more time in group work and less time engaged in strategies that were led from the front than the control group. This is supported by the control group having a greater percentage of observed points where the pupils were engaged in individual work (16.1% compared to 9.2% in the treatment group). This group work element may also supports the observation of 13.3% of points for the treatment group were from pupils researching a question, whereas this was only seen in 1 instance in the control group (0.6%). Although the counts for practical work observed in the control group are half that of the treatment group, 11 to 25, this accounted for a similar percentage of points, 19.6% in the control and 22.9% in treatment, (there were a different number of recorded lesson observations from the two groups) therefore the amount of practical work carried out between the groups was similar.

Therefore, in examining the data using the chi-squared test the differences detailed above are supported by statistical analysis (*Table 4.2*). The chi-squared value is greater than the critical value (4d.f. p>0.05) indicating that there is a less than 0.05 probability that the association was found by chance, therefore there is a statistically significant association between the teaching styles and whether they are used in the control or treatment group, therefore the null hypothesis is rejected. This supports the claim that that the treatment group was receiving a different style of teaching compared to the control group.

The association is very strong even when bearing in mind some method drawbacks (discussed in section 3.14-3.22). This evidence for a difference in the teaching received by the control and treatment groups, permit analysis of the two sets of data, as two different groups from the same population.

4.5. Section B - Comparison of the pupils’ critical thinking test scores before and after being taught the inquiry-based scheme of work

Pupils were tested using the Cornell Critical Thinking Test – Level X (CCTT) to investigate the impact of an inquiry-based scheme of work on their critical thinking skills. This took place in class at the end of the preceding three-topic learning cycle and then at the end of the experimental period (discussed in 3.11)**.** The test was multiple choice and the calculations used to find each pupil’s score are outlined in methodology. Only valid responses have been included; only individuals that completed tests in both pre and post-test events were counted as valid.

4.6 Analysis

Null hypothesis: There is no significant difference between the pupils' CCTT scores pre- and post-experimental period.

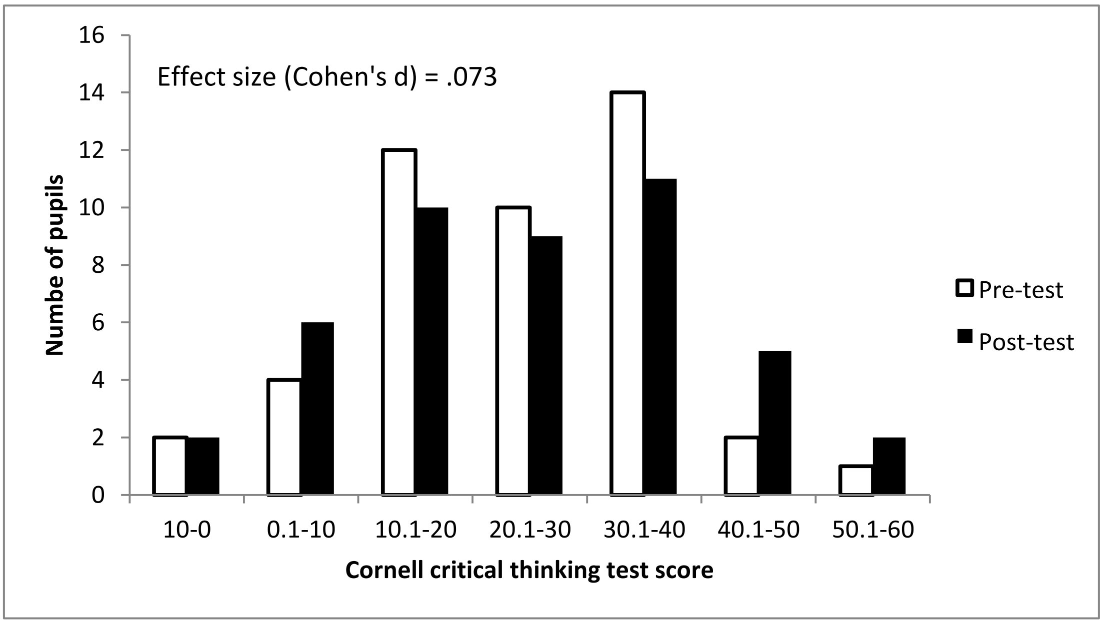
In order to test the hypothesis that there was a significant difference first a test for normality was carried out. There is a normal distribution if the *Shapiro- Wilk* *p* value is higher than 0.05

The data is normally distributed (Appendix 4a – *Table 4.3*) and so to test whether there is a significant difference between the pupils' CCTT scores pre- and post-experimental period then a *paired t-test*[7] is used. When testing for a significant difference there is no statistically significant difference between the pre- and post-scores for CCTT if the *paired t-test* *p* value is greater than 0.05.

Table 4.6 Shows the p value calculated from the paired t-test and the effect size when the control and treatment group, and the treatment group taught by the researcher's pre-and post-tests are compared.

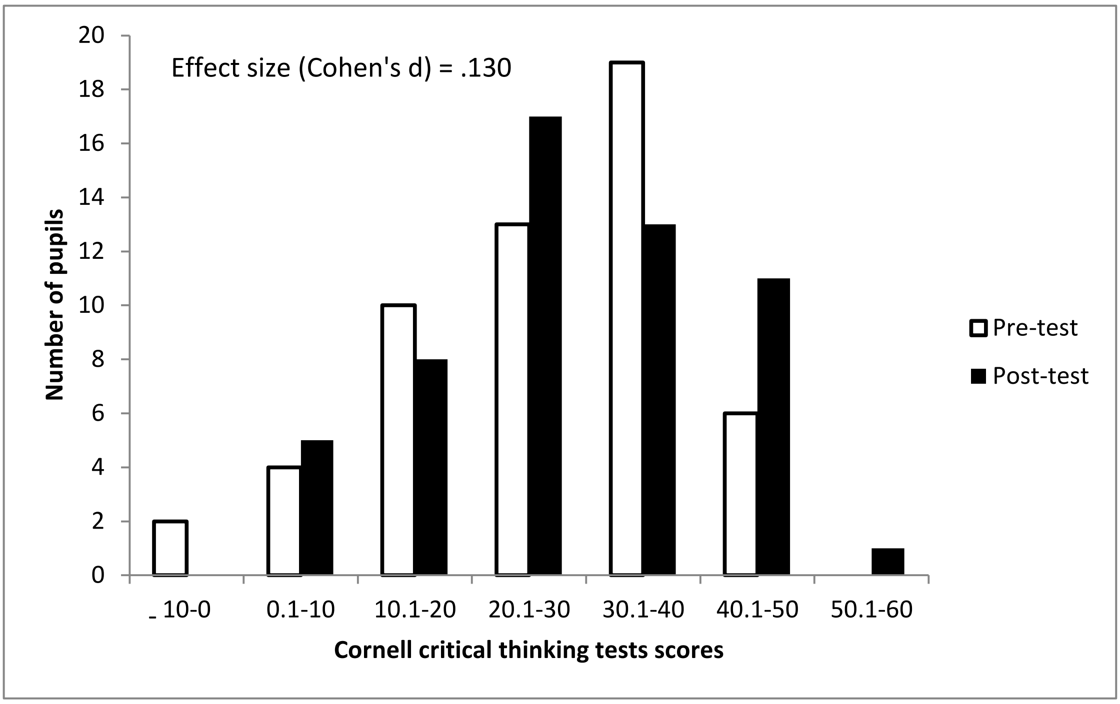
Control or treatment Comparison		Paired Differences					t	df	*p* value	Effect size Cohen's d
		Mean difference	Std. Dev.	Std. Error of diff. in mean	95% Confidence Interval of the Difference					
					Lower	Upper				
C	Pre to post-test score	-.989	9.60	1.43	-3.87	1.90	-.691	44	.493	.073
T	Pre to post-test score	-1.53	8.84	1.19	-3.92	.863	-1.28	54	.209	.130
T researcher-taught)	Pre to post-test score	-2.9	9.22	1.68	-6.38	.511	-1.74	29	.092	.212

Graph 4.1 Comparison of the pre and post-test scores for the control group

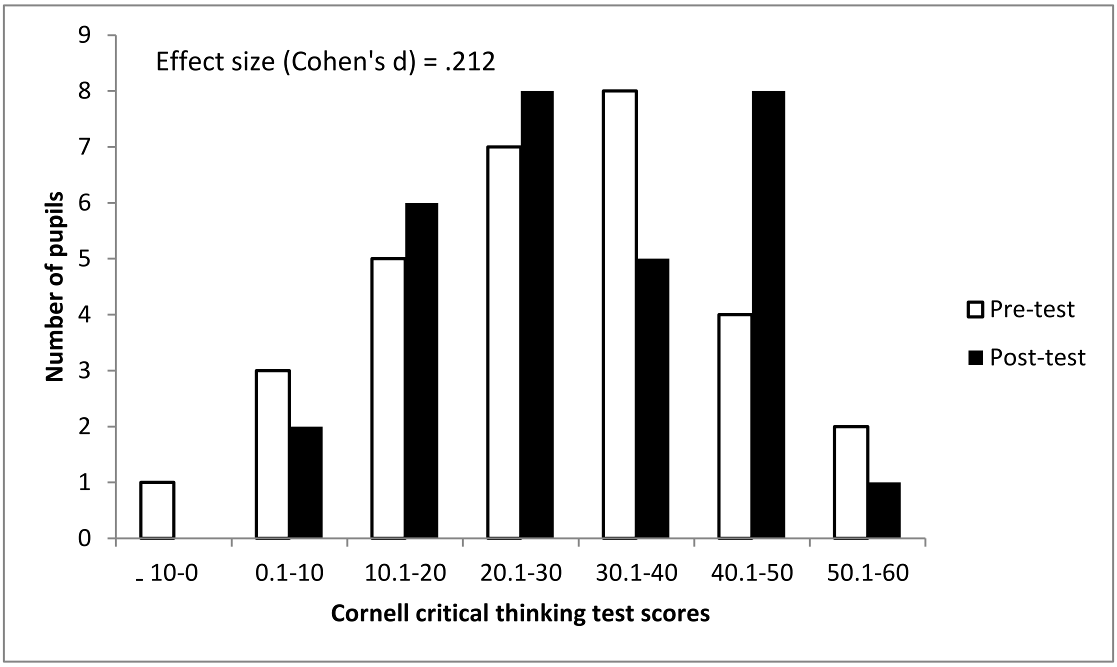


[7] Paired samples which come from the same population and are normally distributed

Graph 4.2 Comparison of the pre- and post-test scores for the treatment group



Graph 4.3 Comparison of the pre and post-test scores for the treatment group taught by the researcher



4.7 Findings

A normal distribution would be expected for mixed ability groups as modelled in this research and particularly as the test was not subject-specific so did not rely on any previous scientific teaching or understanding. In all cases the mean and the test statistic indicate that critical thinking test scores increased over the experimental period (Appendix 4A -*Table 4.3* and *4.4*) and therefore both types of teaching had a positive effect, although the whole treatment group's increase was greater. The researcher's class had almost double the increase (2.9) when compared to the whole treatment group (1.53) and three times greater than the control group (0.989) as shown in *Table 4.6*; however, all of these gains were small when compared to the maximum test score of 70 that could be achieved. Cohen's d (*Table 4.6*) indicates that the effect size when measured by the number of standard deviations difference between the pre- and post-test score is very small for all three groups with the highest value of 0.212 for the researcher's class. This small effect size is also supported by the standard error for the difference in the means (Table 4.6) where the margins for error are similar in value to the differences themselves indicating that changes were not significant.

The effect sizes are represented in *Graphs 4.1-4.3*; for clarity the scores have been categorised into numerical blocks of ten. The improvement in test scores in all three groups can be seen from the pre- to the post-test. In the treatment groups there is a decrease in pupils in the score range -10 to 20, however in the control group this number stays the same. The number of pupils in the higher score bands (40-60) increases in all groups (*Graphs 4.1 – 4.3*) but the increases are greater in the treatment group as a whole and the researcher's class. This suggests that the teaching received by the treatment group had a positive effect on critical thinking skills for all abilities of pupils, whereas the teaching received by the control group had some effect on the critical thinking skills of higher ability pupils but no effect for lower ability pupils.

Examination of the data using a paired-samples t-test (*Table 4.6*) indicates that the prior CCTT scores for the control group (M = 24.4, SD = 13.4) were not statistically significantly different to the CCTT score post experimental period (M = 25.4, SD = 14.1), $t(44) = .69$, $p > 0.05$, Cohen's d = 0.073. Similarly a paired-samples t-test indicated that the prior CCTT scores for the treatment group pre-experimental period (M = 27.1, SD = 12.7) were not significantly different to CCTT score post experimental period (M = 28.7, SD = 12.2), $t(54) = 1.28$, $p > 0.05$, Cohen's d = 0.13. This is also true of the treatment group taught by the researcher (M = 26.7, SD = 13.9) where there was no significant difference between pre- and post- CCTT scores (M = 29.6, SD = 13.5), $t(30) = 1.74$, $p > 0.05$, Cohen's d = .212. Therefore the null hypothesis for control and both treatment groups is accepted.

In response to the research question the analysis indicates that over the period of time measured, the critical thinking skills of pupils did not significantly change, this was true of the control, treatment group and the researcher's class. The standard deviation stayed fairly steady for both groups and this along with the positive correlation between the pre- and post-test scores (*Table 4.5*) supports the validity of the CCTT for consistency in measuring pupils' critical thinking skills.

4.8 Section C - Comparison of the treatment and control groups' change in critical thinking test scores before and after being taught the inquiry-based scheme of work

The findings from the test analysis above demonstrated that there was no significant change in critical thinking test scores for the control, whole treatment group or in the researcher's class. However, a small positive difference can be seen in the mean pre- and post-test score for all groups. As research question one is looking for a comparison between the two groups this next analysis is to determine whether there was a significant difference between the mean change in test score for the treatment group compared to the control group. This is required as the quasi-experimental approach is being used to determine whether inquiry-based teaching has an effect compared to the normal teaching practice in the research school. These findings can be determined through the testing of the null hypothesis stated below. Therefore, this analysis is to determine whether there was a significant difference between the mean change in test score for the treatment group compared to the control group.

4.9 Analysis

Null hypothesis: There is no significant difference between the control and treatment group' change in CCTT scores

In order to test the hypothesis that there was a significant difference first a test for normality was carried out. There is a normal distribution if the *Shapiro- Wilk* *p* value is greater than 0.05.

The data is normally distributed (Appendix 4B – *Table 4.7*) so to test whether there is a statistically significant difference between the change in pupils' CCTT scores in the control and treatment groups an *independent samples t-test*[8] is used. It is not obvious that the difference between pre- and post- test scores would be normally distributed as the data is a change rather than a baseline measure within a class which could be expected to show normal distribution. When testing for a significant difference there is no significant difference between the treatment and control groups' change in CCTT scores if the *p* value is higher than 0.05. The exploratory statistics calculated are shown in Appendix 4B (*Table 4.7*)

[8] Used as there are two sets of parametric data where the means are being compared in two unrelated groups that are from the same population

Table 4.8 Shows the p value calculated from an independent samples t-test when the change in the control and treatment groups' critical thinking test scores are compared.

		Levene's Test		t-test for Equality of Means						
		F	Sig.	t	df	*p* value	Mean Diff.	Std. Error of Diff.	95% Confidence Interval of the Difference Lower	Upper
Change in Tests	Equal variances assumed	.447	.505	-.291	98	.771	-.538	1.85	-4.20	3.13
	Equal variances not assumed			-.289	90.7	.773	-.538	1.86	-4.24	3.16

Table 4.9 Shows the p value calculated from an independent samples t- test when the change in control group and researcher-taught treatment groups' critical thinking test scores are compared

		Levene's Test		t-test for Equality of Means						
		F	Sig.	t	df	*p* value	Mean Diff.	Std. Error Of Diff.	95% Confidence Interval of the Difference Lower	Upper
Change in Tests	Equal variances assumed	.467	.496	-.873	73	.386	-1.94	2.23	-6.38	2.50
	Equal variances not assumed			-.880	64.0	.382	-1.94	2.21	-6.36	2.47

4.10 Findings

Levene's test is used to find whether there is a difference in the variances between the two sets of data. The significance value in this case is .505 (*Table 4.8*) and .496 (*Table 4.9*). As these values are greater than 0.05, the variability in the control and treatment groups test score change is not statistically significantly different; this means the top row of data is used. An independent-samples t-test indicated that there is no significant difference in the change in CCTT scores between the control group (Appendix 4B – *Table 4.7*) (M = 0.99, SD = 9.60) and the whole treatment group (M = 1.53, SD = 8.85), $t(91)$ = .291, $p < 0.05$, d = 0.06, as it can been seen from *Table 4.8* that the p value is greater than 0.05.This is also true of the researcher's class (*Table 4.9*) where (M = 2.93, SD = 9.22), $t(30)$ = .873, $p < 0.05$, d = 0.21.Therefore, the null hypothesis is accepted, indicating that over this test period the groups' mean critical thinking test scores were not statistically affected in either group or when comparing the control to the treatment group.

Therefore, the general finding of research question two would be that that teaching pupils using inquiry-based teaching strategies had no effect on the critical thinking skills. In addition, despite there being greater differences in the exploratory statistics

(*Table 4.7*) when comparing the control group to the researcher's class, the statistical tests indicate that these differences were not significant and therefore in answer to research question 2b, delivery variables did not affect the outcome.

4.11 Section D - Comparison of critical thinking test scores against individual attainment scores

This analysis was carried out to see the correlation between critical thinking skills and attainment (Attainment results in Chapter 5). It provides the opportunity to analyse the interdependence of the ability to attain in subject knowledge-based tests and pupils critical thinking test scores. This will provide the opportunity to discuss whether these measures are two independent skills sets, and therefore can be developed independently in a classroom, or whether they are intrinsically linked and so a rise in one will promote the same in the other.

4.12 Analysis

Null hypothesis: There is no significant correlation between the pre-attainment level or target level and the critical thinking scores of the cohort.

This test is to find correlations therefore, to choose the test it must be established if the data is a) continuous (scaled) and/or ordinal, which the CCTT scores and attainment data are, and b) normally distributed. There is a normal distribution if the *Shapiro- Wilk* *p* value is higher than 0.05.

The target levels and pre-attainment levels are not normally distributed (Appendix 4A – *Table 4.10*), therefore, the test that will be used for two comparisons is *Spearmans Rank correlation co-efficient*[9]. This will give a value between -1 and 1.

Table 4.11 Shows the correlation coefficient calculated from the Spearmans Rank test when comparing all pupils pre-attainment test score to their CCTT score and when comparing their target grade to their CCTT score

	Pre-attainment vs CCTT score	Target grade vs CCTT score
N	86	86
Spearman's Coefficient (*r*)	.64	.50
Critical Value for *p* <0.05	.212	.212
Significant	Yes	Yes

4.13 Findings

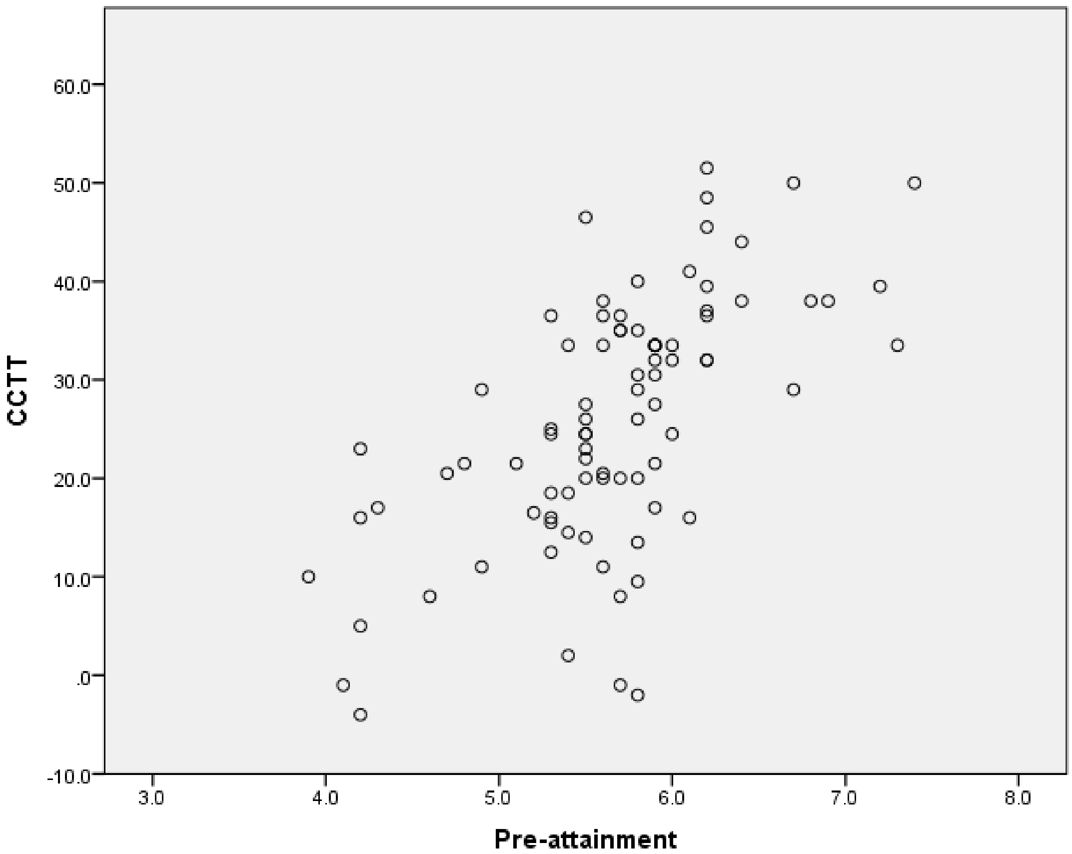
Graph 4.4 shows that pre-attainment scores and CCTT test scores were positively correlated, Spearman's $r(86) = .64$, $p < 0.05$. It is clear that the correlation is not consistent for all groups as the range in critical thinking test scores in pupils with an attainment level between 5 and 6 is very large, ranging between -6 and 42. Pupils

[9] Used to calculate the strength of a correlation between two non-parametric sets of ranked data

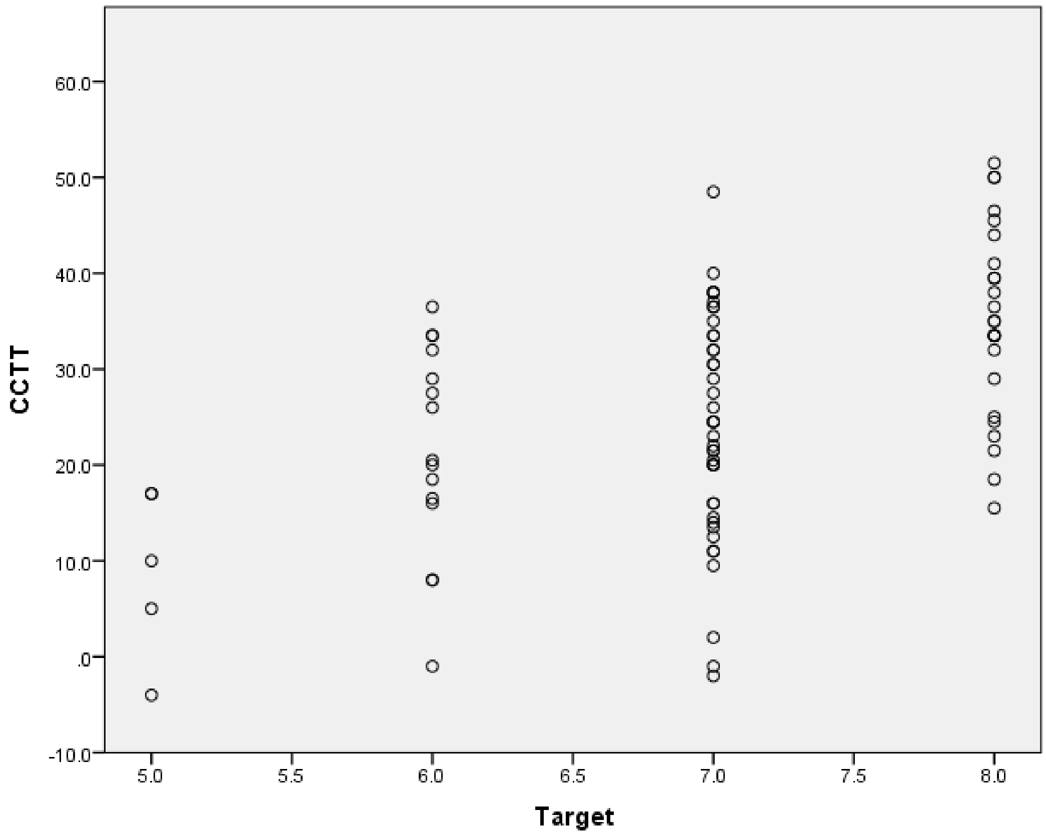
with attainment levels over 6 all achieved critical thinking scores ≥ 30 and pupils with attainment levels less than 5 achieved critical thinking scores ≤ 30. This indicates that for both higher and lower ability pupils their critical thinking scores can be correlated with their attainment indicating the interdependent nature of the two skill sets.

Graph 4.5 shows that target grades for the end of Key Stage 3 and CCTT test scores were positively correlated, Spearman's $r(86) = .50$, $p < 0.05$. This correlation, although significant, is slightly weaker than the CCTT scores and pre-attainment scores. The statistical test was carried out with attainment targets to allow for the fact that the pre-attainment measure in *Graph 4.4* was based on a single test, on a single day which covered a limited number of scientific topics and therefore may not be fully representative of an individual pupil's ability.

Graph 4.4 Comparison of pupils' pre-attainment test scores against their critical thinking test scores



Graph 4.5 Comparison of pupils' target levels against their critical thinking test scores



However, as can be seen from *Graph 4.5*, the correlation is similar as is the differential findings, with there being a clear correlation with pupils with the lowest and highest target grades, but large overlaps in critical thinking test scores for those with target levels of 6 and 7. This could indicate that when considering a target grade, which is derived from an average of English and Maths SATs test scores at the end of Key Stage 2, the interdependence of the two skill sets is not clear. However, it is also the case that target levels are too much of a crude measure and as such incorporate a lot of variability in pupil scoring and so production of a real conclusion is difficult.

4.14 Discussion for RQ1

The first hypothesis tested for this research question indicated that the control and treatment groups increased their mean score for critical thinking skills, however neither of these were statistically significant. It can be seen that the treatment class taught by the researcher showed greater gains, however, this too was statistically insignificant. This was confirmed by the effect size (Cohen's d), although d in the treatment group was almost double that of the control group (0.13 to 0.073) and researcher's class's class was almost three times greater than the control (0.212), these are still small effects. However as an effect was detected a power analysis was carried out on the researcher's class data, indicating that a sample size of 177 would be needed in a two-tailed matched pairs test to show an 80% statistical significance. As the entire cohort was smaller than this, an effect size that shows statistical significance would not have been demonstrable with this study.

As there is no standardised measure of how critical thinking skills progress over a given time period, it is unclear whether the increases, although not significant, are within normal range. Furthermore, any changes in critical thinking (as opposed to attainment) will be affected by more teaching variables: thinking skills develop with life experiences and opportunities that will occur in other subjects. However, this is accounted for by the quasi-experimental study where both conditions are exposed to this variety, making a comparative measure between groups valid.

Although the increases in critical thinking test scores were not significant the increase in the treatment group was greater than the control group; this is in contrast to the results for attainment (Appendix 5A – *Table 5.1*). The correlation found between prior attainment and critical thinking scores (*Graph 4.4)* illustrate a relationship that indicates interdependence in their nature. However, the pattern of change in these scores for the control and treatment groups do not support this finding, as the critical thinking score increase in the treatment group was greater than the control group but the increase in attainment was the opposite (Chapter 5); leading to the suggestion that these skills could develop independently of each other. There may also be other factors affecting pupil performance in critical thinking test; *Graph 4.4* and *4.5* show that, for middle ability pupils, there is more variability. Factors such as inconsistency of motivation levels, confidence or literacy skills may have a greater effect when compared to the lower and higher bands. . It is unlikely that a floor or ceiling effect has occurred as the graphs show no pupils scored in the highest or lowest score brackets in either measure.

In answer to the research sub-question the results indicate that having the researcher teach part of the treatment group did not lead to a delivery variable which had a noticeable effect on the results. The differences that can be seen between the researcher taught class and the treatment group as a whole were not statistically significant.

The second hypothesis was concerned with the difference between the control and treatment groups change in attainment scores, and it was found that there was no statistically significant difference. This answers the research question directly, as any changes in critical thinking scores in the groups being taught by inquiry are compared to a control group in order to measure its effect as the treatment. The teaching of an inquiry-based scheme of work had no impact on pupils' test scores in comparison to pupils' who received the normal teaching style. There was a greater difference in the change between the control group and the researcher's class, in comparison with the entire treatment group.. This could be expected, as the researcher who had developed the scheme of work, had knowledge of the skills that would be assessed in the critical thinking test, and may have inadvertently taught in a way which developed these skills. However this difference was not statistically significant and so in answer to the research sub-question, there was no statistically significant experimenter effect on critical thinking skill tests scores taking place within this study.

Discussion turns to the method and data produced and whether it answers the research question. The researcher did not write the assessment instrument for critical thinking skills, (section 3.11) therefore removing experimenter bias from its design. The Cornell Critical Thinking Test – Series X was developed by academics and has been extensively peer reviewed. Scores produced are calculated from one point for every correct answer and -0.5 for an incorrect, this produces a more valid; total it accounts for pupils making guesses, as the other option is to leave it blank. The scores produced were normally distributed, as expected for a mixed ability group which had no prior teaching in the subject matter being assessed. . The pre- versus post-test scores (*Graphs 4.1 – 4.3*) correlated positively indicating the test's suitability for its purpose of comparing individual pupil scores. The continuous data produced allowed a fine analysis of pre- and post-test scores in the statistical tests and the opportunity to categorise scores more broadly so that changes and effect size could be displayed in graph format.

One drawback of the data analysis was that a large group of pupils were removed due to incomplete the pre- or post-tests. In many cases, these were pupils classed as lower ability (as discussed in section 1.4) who also faced difficulties with their literacy skills; these pupils appeared to take longer to read and understand the test, and this in turn appeared to challenge their motivation towards completing the test. The test was suitable for this year group but at the lower end of the recommended range, which suggests that pupils with a reduced reading age may have struggled to understand and/or complete. Removing this data means that the lower end of the distribution would have a smaller sample size on which to draw conclusions. However, a normal distribution of test scores was found, suggesting that the sample size was large enough to be representative.

The findings in this research are in contrast to those of Ernst and Monroe (2007) and Daud and Husin (2004) where statistically significant increases in critical thinking tests scores were found when using the same critical thinking test in similar quasi-experiments. In the former, although t raw scores were not available for comparison, with a larger sample size, than in this research, and over a one year period, teaching using an environment-based curriculum had a significantly positive effect. A similarly positive trend, although not statistically significant, was seen in the current research but over a shorter time period; perhaps a hypothesis could be developed which would mirror the findings of Ernst and Monroe over a one year period.

Daud and Husin's (2004) reported results can be compared to this research; they found a mean increase of 5.14 marks in the experimental group and 4.85 marks in the treatment , a difference of +0.29 marks attributed to the intervention. This figure is smaller than the change calculated for this research but was found to be statistically significant when ability differences between the control and treatment groups were factored in. This analysis was not conducted for the current research; the sample size was three times bigger than that of Daud and Husin, and therefore representation was assumed, however, this could be an area for development in further research.

In both examples above teaching was not explicitly for promoting critical thinking skills as was the case in research by Miri, David and Uri (2007) and supported by Marin and Halpern (2011). Miri *et al*. used the Californian Critical Thinking Test in a quasi-experiment and found a statistically significant increase in critical thinking skills over a one year period. During this time teachers chose teaching methods to improve higher order thinking skills, and on analysis "Three teaching strategies were identified as promoting higher-order thinking skills… [including] fostering inquiry-orientated experiments." (p.366), this finding supports the objectives of this current research. Again, over a longer time period perhaps the gains reported by Miri *et al.* may be reflected. Further findings reported after a post-post-test (three years) there were no further significant gains suggesting that the explicit instruction had the effect rather than other learning or life experiences. This is supported by Marin and Halpern who assert that "effective critical thinking skills [instruction]…engages students during a period in which a particular skill is introduced, requires deliberate practice." (p. 3); Suggesting that deliberate inclusion of critical thinking skills being assessed into inquiry-based schemes of work, and repeated teaching of these may have led to greater gains.

This idea is of interest within the context of the research school, which teaches the National Curriculum, as it is stated in the handbook (cited DCSF, 2008) "The curriculum should enable pupils to think creatively and critically to solve problems…by providing rich and varied contexts" (p. 3) There is no instruction for these skills to be taught explicitly it is however suggested (DCSF, 2008) that their definition of critical thinking skills is in line with this research, as they too quote from Ennis (1987);

suggesting that developing teaching strategies to develop critical thinking skills is in concordance with the national agenda.

In terms of the national agenda, the results from PISA are given significant weight. The analysis carried out by McConney *et al.* (2014) of PISA (2006) suggest that although inquiry-based learning is measured to have a negative effect on students' performance, they argue that these tests do not account for other skills that may have been developed alongside or instead of those measured such as depth of understanding or the positive effects inquiry can have on thinking skills such as drawing conclusions and evaluating data, both of which are assessed as part of the critical thinking test. However, in contrast to the assertions of McConney *et al.,* these skills were not seen to be improved in this study. This is also true of the cognitive domains researched by Furtak *et al*.(2012); although the questions in the critical thinking test could be sub-divided into categories which may indicate epistemic or procedural improvements, the sample size was too small to draw any valid conclusions from such an analysis and as such, this was advised against by the test designers and reviewers. It is also true that while Lazonder and Harmsen (2016) found that "guidance has a larger impact on the development of inquiry skills than on the acquisition of domain knowledge" (p. 23), as the skills being using for inquiry in this study were being tested through critical thinking then their findings were counter to those of the current research where no improvements were found.

The choice of critical thinking test was discussed in methods (section 3.11), and has precedents in this type of research; however, the test did have drawbacks. As there are few tests written for the age group used in this research and the cohort were at the lower end of the range that the test used was designed for. The test only assessed four areas of critical thinking; induction, deduction, credibility of assertions and identifying assumptions, and although the total mean score for these skills was taken to be a valid indicator of critical thinking overall, this list is not comprehensive. More emphasis on the critical thinking skills in UK education may increase the development of tests appropriate to all age groups. The accessibility of these tests may encourage an increased teaching of these skills, so that any gains can be measured within the school context. In summary of the findings for research question 1 'What is the impact of an inquiry-based scheme of work on pupils' critical thinking skills?', it was found that although the teaching style did have a positive impact on critical thinking skills in both groups, the increases were not statistically significant. This is also true when comparing the change in critical thinking scores for control and treatment groups; although there was a greater change in the treatment group, and greater again for the researcher's class, however no differences were statistically significant. Therefore, teaching using inquiry-based strategies had no statistically significant impact on pupils' critical thinking skills. In the research sub-question 1a 'Is there a delivery variable that influences these effects?', it was found that having the researcher teach one of the treatment classes similarly had no statistically significant effect on the pupil scores.

Chapter 5 - Data analysis, findings and discussion

5.1 This chapter will present the data for research question 2 and 2a as detailed below; it includes findings, analysis and discussion sections. It comprises of two analyses of attainment score comparisons; the first comparing pre-test to post-test scores and the second comparing the change in those scores between the control and the treatment group.

Research Question 2 – What is the impact of an inquiry-based scheme of work on pupils' attainment?

Research Sub-question 2a – Is there a delivery variable that influences these effects?

In analysing these data the probability (p) value that all of the statistical findings will be measured against is 0.05. If $p \geq 0.05$ then there is an equal or greater than 5% probability that the findings are due to chance and therefore the null hypothesis cannot be rejected.

5.2 **Section A - Comparison of the pupils' attainment scores before and after being taught the inquiry-based scheme of work**

Pupils were tested using a departmental standardised test for scientific content understanding to investigate the impact of an inquiry-based scheme of work on their level of attainment. This took place in class at the end of the preceding three-topic learning cycle and then at the end of the experimental period (discussed in section 3.10). The results were divided into ten sub-levels, in order that the raw scores form the two tiers of papers could be compared, and therefore pupil scores were recorded to one decimal place. Only valid responses have been included thereby only individuals who completed tests in both pre and post-test events were counted as valid.

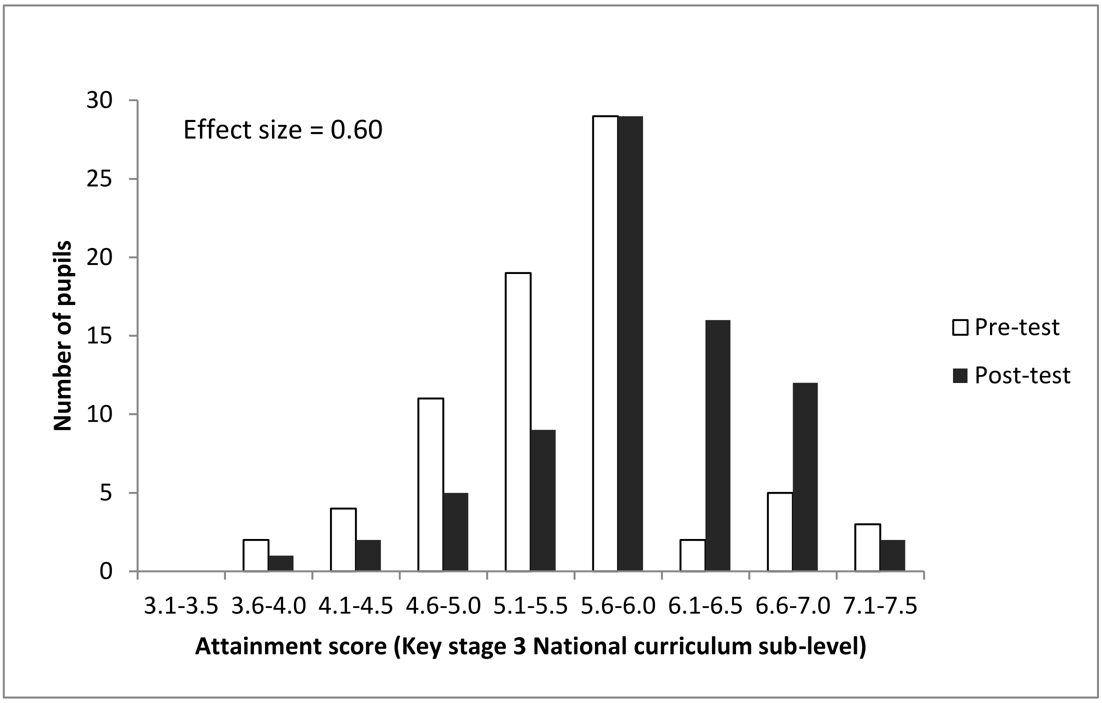
5.3 **Analysis**

Null hypothesis: There is no significant difference between the pupils' attainment scores pre- and post-experimental period.

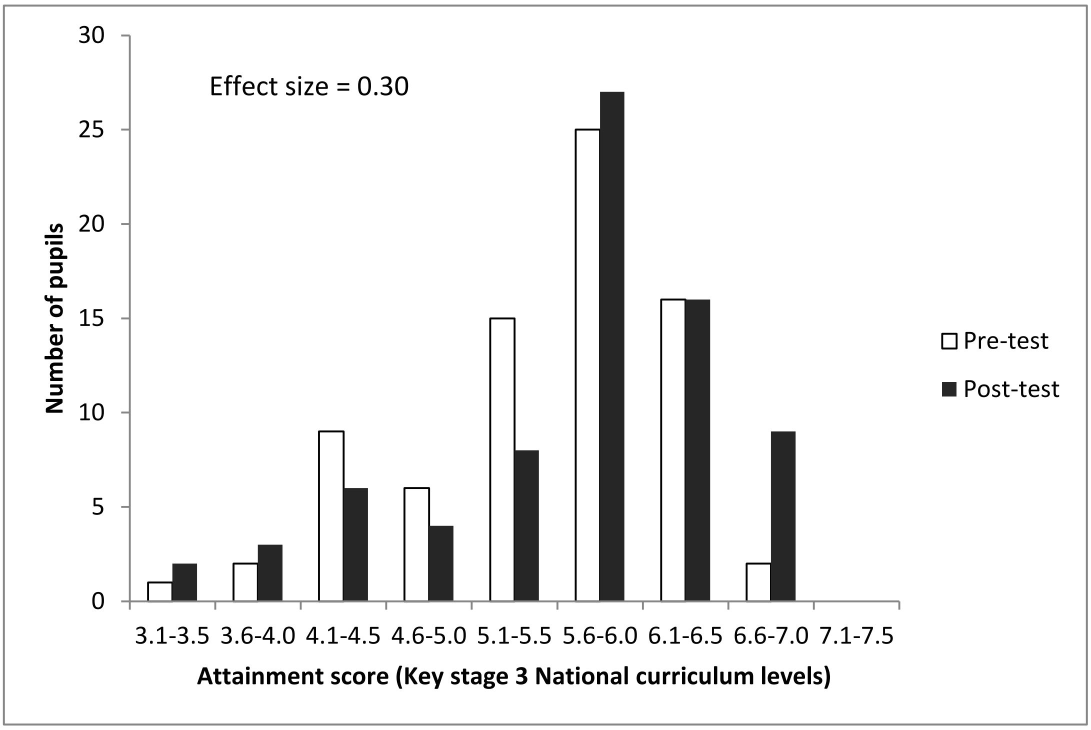
In order to test the hypothesis that there was a significant difference first a test for normality was carried out. There is a normal distribution if the *Shapiro- Wilk* *p* value is higher than 0.05.

Some of the data is not normally distributed (Appendix 5A -Table 5.1) and so to test whether the two populations of data are the same between the pupils' attainment scores pre- and post- scheme of work then a *Wilcoxon signed ranks test*[10] is used. When testing for a significant difference there is no statistically significant difference between the pre- and post-test scores for attainment if the *Wilcoxon signed ranks test* *p* value is greater than 0.05.

Graph 5.1 Comparison of the pre and post-test scores for the control group



Graph 5.2 Comparison of the pre- and post-test scores for the treatment group



[10] Paired samples which come from the same population but are not normally distributed

Graph 5.3 Comparison of the pre and post-test scores for the researcher's class

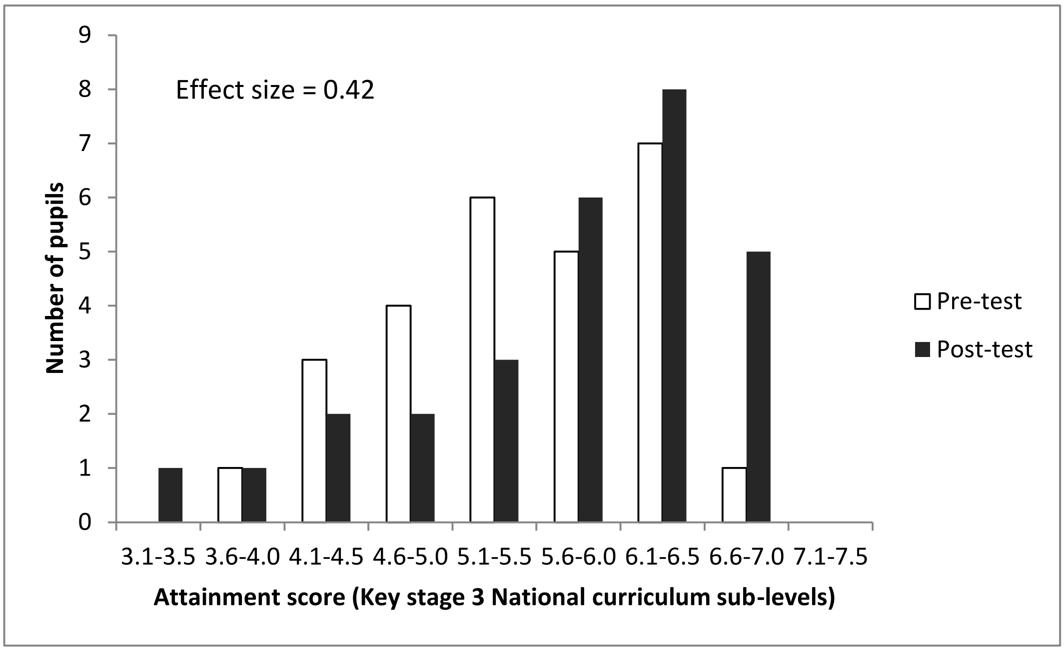


Table 5.3 Shows the p value calculated from the Wilcoxon signed ranks test and the effect size when the control and treatment group, and the researcher's class's pre- and post-tests are compared.

Control or treatment	Test Statistic calculated	Test statistic value	Effect size	Pearsons correlation coefficient (r)
Control	Z (test stat) Asymp. Sig. (2-tailed) *p* value	-5.20 .000	-0.60	0.60
Treatment	Z (test stat) Asymp. Sig. (2- tailed) *p* value	-2.62 .009	-0.30	0.30
Treatment taught by the researcher	Z (test stat) Asymp. Sig. (2- tailed) *p* value	-2.16 .031	-0.42	0.77

5.4 Findings

As most of the data is not normally distributed and the findings need comparable statistics the mean is replaced by the median as it will better represent the centre of the distribution. This is supported by the anecdotal evidence that the intake of pupils in the research school is known to have a general skew towards higher levels. In addition, sub-levels were used rather than raw data, (discussed in section 3.8) which would also make normality more difficult to demonstrate. The median and the test statistic indicate that both the control and treatment groups increased their attainment scores over the experimental period (Appendix 5A *Table 5.1*) and therefore both types of teaching had a positive effect, although the control groups median increased more, 0.30 compared to 0.10 for the whole treatment group. There was a consistent variance and range in the control group (Appendix 5A *Table 5.1*) pre- and post-test compared to the whole treatment group and the researcher's class. This indicates that differences between the test scores and the mean test score were consistent; this implies that the control group's data set was more stable from the beginning to end of the experimental period than the treatment group.

The effect sizes stated in *Table 5.3* are represented in *Graphs 5.1-5.3*, for clarity the scores have been categorised into 0.5 level groups. The improvement in attainment scores is clear in all three groups. There is an increase in pupils receiving the lowest levels in the treatment group as a whole and in the researcher's class which is not reflective of the control group. However the increase in pupils achieving level 7+ increased in all of the groups with a slight decrease in the 7-7.5 category for the control group. This suggests an uneven effect of the intervention on different parts of the distribution.

Examination of the data using a Wilcoxon Signed-ranks test indicates that the prior attainment score for the control group (*Mdn*=5.60) was statistically significantly different to the attainment score post experimental period (*Mdn* = 5.90), Z =5.20, $p < 0.05$, *r (Pearsons correlation)* = 0.60. Similarly a Wilcoxon Signed-ranks test indicated that the prior attainment score for the treatment group (*Mdn* =5.75) was significantly different to the attainment score post experimental period (*Mdn* = 5.85), Z = 2.62, $p < 0.05$, r = 0.30. This is also true of the researcher's class (Mdn =5.60) was statistically significantly larger (Mdn = 5.90), z = 2.16, p<0.05, r =0.77. Therefore the null hypothesis for control and both treatment groups is rejected.

The Pearsons Correlation[11] (*Table 5.3*) indicates that for both groups and the researcher's class treatment group alone there is a positive correlation between pre- and post-test scores, supporting the validity of this test as a good measure of pupils' attainment in this setting. The correlation is larger in the control group than the treatment group which is consistent with the higher levels of variance seen between the pre- and post-test scores in the treatment group. However the most positive correlation (r = 0.77) is in the treatment group that was taught by the researcher.

[11] Comparison of two continuous variables to find the strength of the linear relationship

5.5 Section B - Comparison of the treatment and control groups' change in attainment test scores before and after being taught the inquiry-based scheme of work

The findings of the test analysis above demonstrated that there was a significant change in both the control and treatment groups' attainment test scores. However, as research question two is looking for a comparison between the two groups this next analysis is to determine whether there was a significant difference between the mean change in test score for the treatment group compared to the control group. This is required as the quasi-experimental approach is being used to determine whether inquiry-based teaching has an effect compared to the normal teaching practice in the research school. These findings can be determined through the testing of the null hypothesis stated below.

5.6 Analysis

Null hypothesis: There is no significant difference between the control and treatment groups change in attainment scores

In order to test the hypothesis that there was a significant difference first a test for normality was carried out. There is a normal distribution if the *Shapiro- Wilk* *p* value is greater than 0.05.

Not all the data is normally distributed (Appendix 5B - *Table 5.4*) so to test whether there is a statistically significant difference between the change in pupils' attainment scores in the control and treatment groups a *Mann- Whitney U test*[12] is used. When testing for a significant difference, there is no significant difference between the treatment and control groups change in attainment scores if the *p* value is higher than 0.05. The ranks calculated in the statistical test are shown in Appendix 5B (*Table 5.5 and 5.7*)

Table 5.4 Shows the test statistic calculated from the Mann-Whitney U test when the change in control and treatment groups' attainment scores are compared

Test statistic calculated	Test statistic values
Mann–Whitney U	2250
Wilcoxon W	5331
Z	-2.47
Asymp.Sig. (2-tailed) (*p* value)	.013
Effect size	-.199

[12] Used as there are two sets of non-parametric data where two independent groups are from the same population and have produced continuous data.

Table 5.6 Shows the test statistic calculated from the Mann-Whitney U test when the change in control group and researcher-taught treatment group attainment scores are compared

Test Statistic calculated	Test statistic values
Mann–Whitney U	939
Wilcoxon W	1317
Z	-.559
Asymp.Sig. (2-tailed) (*p* value)	.576
Effect size	-.055

5.7 Findings

A Mann-Whitney U test indicated that the change in attainment for the control group pre- to post-experimental period (*Mdn* = 0.30) was statistically significantly different than for the treatment group (*Mdn* = 0.20), $U = 2250$, $p = .013$, as the *p* value is less the 0.05, this indicates a greater than 0.95 probability that the difference in the distributions was not due to chance. Therefore the null hypothesis is rejected, as the control group's score increased to a median value statistically significantly higher than the treatment group.

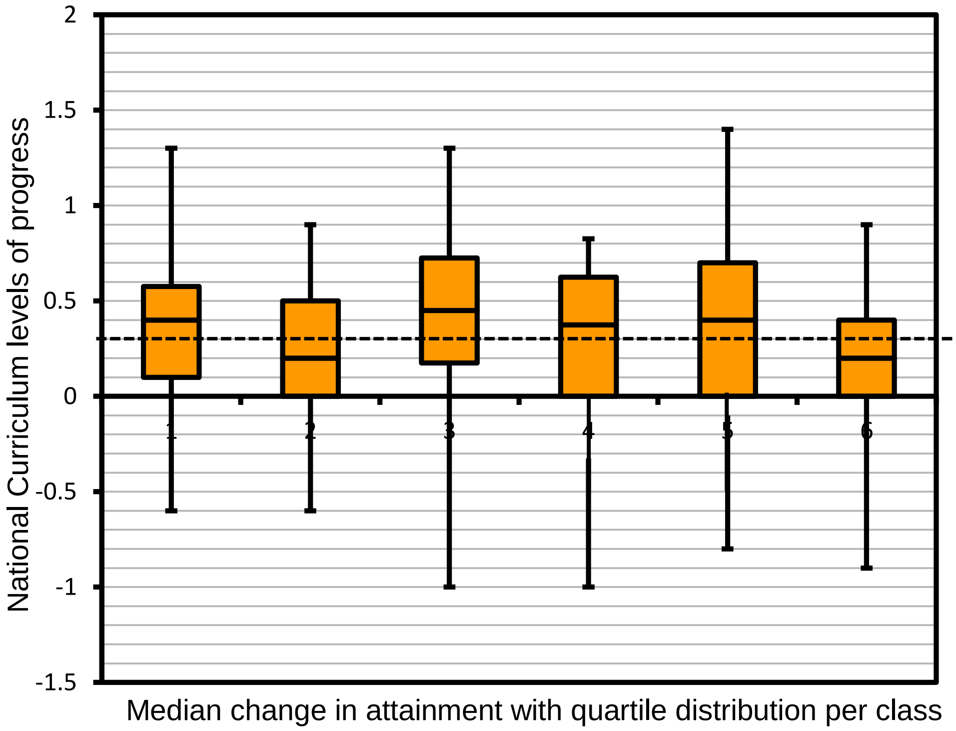
However when comparing the treatment class that was taught by the researcher the median (Mdn = 0.30) is the same as the control group and as U = 939 and $p = 0.576$, and as the *p* value is greater than 0.05 there is no significant statistical difference between the control and the treatment group that was taught by the researcher, demonstrating no statistically significant difference in the increase in attainment scores between the two groups. This is also supported by the effect size (*Table 5.6*) which indicates that the average change in the control group would exceed 50% of the same measure in the researcher's class. This indicates that teaching pupils using the inquiry-based scheme of work had a negative effect on their scientific content attainment progress compared to their usual teaching practice (except in the treatment class taught by the researcher) as the median value for attainment change in the treatment group was shown to be significantly lower than that for the control group 0.20 compared to 0.30 (*Table 5.4*). Therefore, the general finding of this research question would be that there was an effect on pupils' attainment scores when taught three units of work using inquiry-based strategies and that this effect was negative (except in the class taught by the researcher) compared with being taught by a teachers' normal approach.

5.8 Discussion for RQ2

The pre post-attainment tests used (section 3.18) comprised of 50% questions from the three topics in the preceding cycle of learning and 50% all topics previously taught during Key Stage 3. The first hypothesis tested indicated that the median score for both the control and treatment groups increased and in both cases this increase was statistically significant. The increases of 0.10 and 0.30 were based on KS3 curriculum levels with the range achievable being 3 to 7.9 and therefore the small changes represent a fairly large percentage of the total score.

An increase in median of 0.30 is comparable to a KS3 National Curriculum increase of one sub-level, and a 0.10 increase, one third of a sub-level. Over a three year programme of science pupils are expected to make two levels of progress or six sub-levels. Although learning is rarely linear, it is treated as such in this school where pupils receive reported sub-levels three times a year and a lack of progress is an issue that needs to be addressed by practitioners; therefore, there is an expectation of two sub-levels per year (0.67 of a level). The three unit scheme of work taught for this research lasted 8-10 weeks (24 lessons), equivalent to 28% of the school year; therefore, the expected increase in attainment would be 0.18 National Curriculum levels. This data shows that the control group and the researcher's treatment class exceeded this whereas the entire treatment group did not.

Graph 5.4 Shows the median change in attainment for each class with quartile distribution. The expected mean attainment change indicated with the dashed line.



(1-3 =Control classes, 4-6 = Treatment classes, 5 = treatment class taught by the researcher)

When comparing the medians and quartile distributions in *Graph 5.4* there is variation between the classes within the treatment and control groups, with the highest score being achieved in the researcher's class. The treatment and control group each had one of the two lowest medians and upper quartile distributions, this is also true of the lowest change in scores of -1.0. The highest 4^{th} quartile minimum change in score was in two of the control classes at -0.6. This graph suggests that, although the impact of inquiry-based learning seems to be negative, there is wide variation in pupils' attainment progress within the control and treatment groups and therefore other factors may well be affecting this measure.

To answer the research sub-question the results indicate that having the researcher teach part of the treatment group did lead to a delivery variable which had an effect on the results. There was a statistically significant increase in the median (0.4); more than double that of the expected levels of progress over this period. This suggests that when an inquiry-based scheme of work was delivered by teachers other than the researcher, the effect on pupil attainment was negative. However, the lowest value for attainment change in the researcher's class is still lower than in the control classes, indicating that researcher involvement did not completely negate the negative impact of the intervention.

The next hypothesis tested found that the changes in attainment scores were statistically significantly different between the control and treatment group. The teaching of a three unit inquiry-based scheme of work had a negative impact on progress in comparison to the normal teaching style pupils received, except for researcher's class, where pupil progress was not statistically significantly different from the control group. These differences in findings between the treatment group as a whole and the researcher's class suggest, in answer to the research sub-question, that an experimenter effect was taking place within this study, alongside other variables could have conceivably contributed to a difference in the teaching and learning experience of pupils between the six classes.

As the researcher had developed the scheme of work based on extensive reading and literature review, there is good reason to believe that the understanding of both the philosophy and the teaching strategies used was deeper. This scheme of work was given to the teachers who were teaching the other two treatment classes; these included detailed lesson plans, side notes and resources photocopied in advance. This process was to maintain consistency across the treatment groups. The drawback was, that whilst the researcher was very prepared, , the other teachers were less vested in the development, structure or timings of the lessons, so this may have affected their competency and commitment to high quality delivery. Although, acknowledging this as a possible reason for the disparity in attainment outcomes, it should be noted that it is standard practice in schools for teachers to write and resource schemes of work to be followed by other teachers; the level to which this is insisted on varies between schools.

There is also much literature on how teachers' belief systems will affect their delivery of an inquiry-based scheme of work. As stated by Wallace and Kang (2004) "Decisions teachers make in the classroom are based on their own belief system" (p. 938), therefore in spite of teachers being provided with resources, the way that they delivered them may have been affected, as Wallace and Kang continue to assert in this paper that "teachers' beliefs of the capabilities of pupils to deal with inquiry-based learning has a hugely negative effect on its use" (p. 940). This belief system was also commented on by Brown and Melear (2006) who suggested that this was contributed to by teachers' familiarity with this style of teaching and their confidence in delivering it. In both of these studies, alongside Crawford (2000), training for staff was suggested as a major contributor to the success of inquiry-based teaching. The five minute tallies from lesson recordings and pupil feedback from the focus group suggested that teachers did use the resources and delivered the lessons in line with the scheme of work, however the nuances of teaching cannot be ascertained from these e.g. questioning style and scaffolding of understanding.

Furthermore, classes in the treatment group had changes to teachers during the year due to maternity cover and two re-timetabling events; this affected two of the treatment classes, neither were the researcher's class.

The researcher did not write or see the assessment instrument for attainment prior to the class completing therefore removing a source of bias. It was requested that no classes did 'revision' prior to the attainment test, however in the course of 'normal' teaching (as continued in the control group) it would not be unusual for exam questions to be shown or discussed, which would give these groups an advantage over the treatment group.

Alongside these variables, others may have affected this outcome, such as teacher competences in classroom and behaviour management, questioning, organisation and feedback. However, all of these causes of variation are what would normally be expected in a standard secondary school, and highlight the difficulties of attempting to implement a pedagogical innovation in a real-school situation. Therefore, from the results of this study it is not possible, definitively, to attribute the causation of the differing levels of attainment.

Discussion now turns to whether the data produced and the method used can be used to answer the research question. As described in the methods chapter, the attainment test used was consistent with the tests that the research school would normally use to measure pupil progress and therefore is reflective of the school context. The test measured both recall and understanding of scientific facts and was constructed from past SATs levelled questions; moving through the levels, questions incorporated recall, describing, explaining and finally application of knowledge. This method of measuring attainment has been used by other researchers. In the UK, Mant, Wilson and Coates (2007) used a levelled test to find that 10% more pupils had achieved a L5 in SATs "…[after] increasing conceptual challenge in primary science lessons."(p.1707). In the US, Geier, Blumenfeld, Marx, Krajcik, Fishman, Soloway, and Clay-Chambers (2008) used state-wide standardized tests and found an increase in test scores after 7th and 8th graders were engaged in an inquiry-based learning project. Although other research has been conducted into the increasing attainment of pupils engaged in inquiry-based teaching, tests are described as "science content assessment" (Lynch, Kuipers, Pyke, &, Szesze, 2005, p. 926), "criterion referenced tests" (Minner, Century, &, Levy, 2010, p.19) and "applied physics test" (White & Frederiksen, 1998, p.60). Therefore, even though the aim and measure of these pieces of research is the same, the rationale for the choices of assessment instruments are not as closely linked with this research project as those used by Mant *et al*. and Geier *et al*., and therefore, it is to their research that this research can be more closely aligned and compared.

Further support for the use of these attainment tests is the level of correlation shown between the pre- and post-test scores. In both the treatment and control groups there was a positive correlation. This suggests validity in the use of this instrument; as a lack of correlation would indicate that the test was not delivering the information required to answer the question. The correlation was stronger in the control group than the treatment, which may suggest that the change in teaching style for the treatment group affected pupils unevenly , so that they did not all perform as would be expected.

More simply, fewer pupils in the entire treatment group, made the expected linear progression that was modelled by the control group. An inquiry-based scheme of work would need to be, at least, as effective in promoting attainment gains as that expected by 'normal' teaching, in order that it could be used in the school context. A possible reason for this disparity is that the teaching strategies employed were less focussed on teaching to the test, as is the case with normal teaching; inconsistency in teaching style may have led to an inconsistency in pupil progression. This was not the case with the researcher's class where the strongest positive correlation was seen. This supports the discussion above where the researcher acting as a teacher is comfortable with the concepts and resources to make the lessons successful. It may also be the case that the researcher will have used some of the strategies or teaching styles before and as such the pupils may not notice such a change. It could also be argued that the teacher was more adept at using inquiry-based strategies in the service of what pupils needed to know, in terms of subject knowledge, to still prepare them for the attainment test. If this is the case it would explain why the correlation would be stronger for the control group and the researcher's class as the correlation indicates a level of consistency in learning.

It can be seen from the data that the numbers in the analysis were smaller than the number in the control and treatment cohorts; 153 pupils pre- and post-test scores compared to the 176 pupils in the cohort. This was because pupils who were absent for either the post or pre-test could not be included. Absence is often associated with low attainment (DfE, 2015) and/or progress and therefore, data removed may have had upwards effect on the median progress achieved in the two groups. This may be a source of error in conclusions drawn for attainment per group but not on the comparison between the control and treatment groups and therefore the ability to answer the research question; as the assumption, in the absence of data, is that there is no difference in attendance rate between the two groups.

The justification for use of the levelled test have already been outlined in the methods chapter and are summed up by Mant *et al*. (2007) stating that although "Teachers are being encouraged to be creative and to take risks in their teaching…schools are judged by prescribed outcome measures: national test results" (p. 1708). In the same way regular class tests reflect this format and are being used by the schools to hold teachers to account for pupil progress. The attainment test used in this research reflects how attainment is measured and reported nationally, however this does not mean that it is the truest way to measure attainment in an academic sense. This statement has generated discussion in recent UK educational reports (NESTA, 2005; Wellcome Trust, 2011; SCORE, 2008) where, on commenting on the UK assessment framework, they all agree that current instruments do not test the complexity of scientific understanding summarised by Horner (2011),"These methods of assessment do not therefore reflect the integration of knowledge and inquiry upon which science relies." p.14). Therefore, an alternative method for assessing this deeper understanding needs to be explored (section 3.18).

If this is true then, although the attainment of the pupils in the treatment group was measured using a valid instrument, it may have been that pupils 'understanding of scientific ideas' altered in a way that was not measured by this test. A less controversial assertion might be that pupils who were taught by the teachers' normal methods would have been more likely to show attainment gains when measured by this style of test, as over the years teachers strategies involved with 'teaching to the test' have become embedded. However, the research question investigated the effect of an inquiry-based scheme of work on attainment compared to normal teaching strategies, and as such, it must be accepted that any change to the status quo of teaching would not necessarily show its advantage over this norm. The attainment measure used is defined in the introduction and links to KS3 curriculum levels, and therefore the test used does satisfy the question posed.

In the research literature it was found in many cases that attainment was improved by the use of inquiry-based strategies. Geier *et al*. (2008) and Mant *et al*. (2007) both found attainment increases in standardized tests. This was not found in this research. The major themes for researchers and practitioners to take from this and other similar studies are attainment testing and staff training. As already highlighted in this discussion, the effect on pupil attainment from teachers compared to the researcher teaching was statistically significantly different and may have been due to a lack of training. In the research by Geier *et al*. teachers involved in the delivery of the inquiry-based scheme of work had a summer school and in-service support. Mant *et al*. designed the study so that teachers wrote the lessons collaboratively and evaluated them together. Although Geier *et al*. "…do not claim….that inquiry science alone will enhance achievement" (p. 935) they did state that consistency of resources across groups gave greater gains compared to other studies, which supports the design used in this research. In terms of attainment, , alternatives do exist and are being used in research, but as researchers are creating their own content tests comparisons are difficult. Also, as already discussed some are moving to using the instruments such as the flow-map (section 3.18), further research into this area would be of interest to practitioners when assessing pupils in formative ways for their level of understanding.

The finding that attainment was negatively affected is in line with that found in McConney *et al*.'s review of PISA (2006) where pupils who reported high levels of inquiry-based learning achieved lower science literacy scores, which are comparable to the focus of the attainment test measure in this study. This is supported by the meta-analyses by both Furtak *et al.* (2012) and Lazonder and Harmsen (2016) who found that inquiry-based learning did not show improvements in the "conceptual structures" (Furtak *et al*., 2012, p. 305) even with teacher-led inquiry, where the effect size was the greatest, or in "learning outcomes" (Lazonder and Harmsen, 2016, p. 22) irrespective of the level of guidance provided. These two measures were the closest that could be aligned with the attainment test in this research piece and show analogous results.

To summarise the findings of research question 2; ‘What is the impact of an inquiry-based scheme of work on pupils’ attainment?’ It was found that although the intervention did have a positive effect on attainment, the increase was less than that found for the control group. The attainment change for both groups was significant, with the control groups’ change being significantly greater than the treatment groups’, therefore on comparison, the inquiry-based teaching strategies had a negative impact on pupil attainment. In the research sub-question 2a; ‘Is there a delivery variable that influences these effects?’ it was found that having the researcher teach one of the treatment classes did have an effect on attainment that differed from the treatment group as a whole. Positive attainment gains were found in this class equal to those of the control group; this suggests that this delivery variable (section 5.7.3) did have an influence on the ability to measure the impact of inquiry-based teaching strategies on attainment.

Chapter 6 - Qualitative data analysis on teacher and pupil perceptions

6.1 This chapter will present the data for the third research question and include findings, analysis and discussion. The raw data is comprised of three transcripts; a control and treatment pupil focus group and a teacher focus group. Comparisons have been made between the control and treatment groups' responses as well as between the themes discussed by teachers and treatment pupils. Discussion makes reference to quantitative data from chapters four and five to provide some triangulation which will increase the validity of conclusions drawn.

Research Question 3 – What are pupils and teachers perceptions on the use of inquiry-based teaching strategies in science?

6.2 Section A - Comparison of the themes discussed between the treatment and control pupil focus groups

The focus groups were comprised of volunteers taken from the entire cohort who gained permission from their parents (discussed section 3.12). As only nine pupils returned these, accounting for about 5% of the cohort, it was decided to include them all. Three of these were from the control group and six from the treatment group. The pupils were audio recorded as they each responded to very question and then their comments were transcribed (See Appendix A and B).

6.3 Analysis

Table 6.1 (See Appendix 6F) summarises the groups' responses to each question and gives a simple comparison of the answers between the control and treatment groups. *Table 6.2* is a quantitative analysis of the themes mentioned by the two groups of pupils and whether they were in a positive, negative or neutral context. The thematic comparison was produced by colour coding the transcripts (See Appendix A and B) to identify categories and then themes which led on to a semi-quantitative analysis of frequency alongside identification of supporting quotes (See Appendix C). *Table 6.2* shows the actual number of responses first followed, in brackets, with the number of responses that equates to per pupil (frequency) calculated due to the sample sizes being different.

Table 6.2 Quantitative Comparison of the Themes Discussed by the Treatment and Control Groups' Pupil Focus Groups

Theme	Control group (3 pupils)			Treatment group (7 pupils)		
	Positive *(Per pupil)*	Negative *(Per pupil)*	Neutral *(Per pupil)*	Positive *(Per pupil)*	Negative *(Per pupil)*	Neutral *(Per pupil)*
Topic/subject content	6 *(2)*	3 *(1)*	0 *(0)*	5 *(0.8)*	2 *(0.3)*	0 *(0)*
Practical work	5 *(1.7)*	1 *(0.3)*	3 *(1)*	7 *(1.2)*	4 *(0.7)*	0 *(0)*
Teaching strategies	15 *(5)*	1 *(0.3)*	10 *(3.3)*	10 *(1.7)*	17 *(2.8)*	1 *(0.2)*
▫ Group work	1 *(0.3)*	0 *(0)*	0 *(0)*	*3 (1)*	11 *(1.8)*	1 *(0.3)*
Focus/work completion	7 *(2.3)*	3 *(1)*	0 *(0)*	12 *(2)*	7 *(1.2)*	3 *(0.5)*

6.4 Findings for pupil perceptions

Changes to teaching approach

There was a difference in response to question 1 (*Table 6.1* Appendix 6F) between the two groups as in the treatment group 87% of pupils noticed a change in teaching style whereas for the control group the figure was 33%. The control group had changes in teaching staff as stated by a pupil (see Appendix 6A), this may account for one pupil identifying a change.

Group work

There were repeated mentions of group work (15 mentions) by the treatment group, which indicates that this was the largest change that they noticed. It equated to a frequency of 2.5 compared to 0.3 in the control group; and of these mentions the frequency of them being in a negative context was 1.8, suggesting that the major change identified was a negative one. Pupils refer to group work in all six questions, further indicating its importance, although in question 1 three pupils (50%) stated that they enjoyed the change, working together and discussing their ideas, but then went on in the proceeding questions to discuss its drawbacks. Pupils also repeated that working in a group meant that some pupils did all the work whilst others did nothing, this is supported by data in *Table 6.2* and the transcript (See Appendix B) as each comment that was positive (frequency 2) about their own focus there typically followed a secondary comment indicating the lack of focus (frequency 1.2) in other pupils. This is also supported by the ratio of positive to negative comments about pupil's own focus and work completion being similar between the control and the treatment group 1: 0.43 and 1:0.60 respectively.

Practical work

In question 2 (*Table 6.1* Appendix 6F), 100% of the control group said that they had enjoyed practical work in the three units of work whereas 50% of the treatment group said that they had. Although of the references to practical work by the treatment group throughout (*Table 6.2*) the 1:0.58 ratio of positive to negative comments indicates that pupils were more positive about this than indicated in the first question. However, compared to the control group where the ratio was 1:0.2 there is a large difference. This is supported by pupils' comments to question 2 in the treatment group (*Table 6.1* Appendix 6F) where they felt that working in small groups and having to help others as well as having to plan their own investigations meant that they felt rushed, although in question 4 only one pupil (17%) felt that planning investigations and working in groups affected her ability to learn.

Topic content

There was more than double the total number of mentions for enjoyment of the subject matter in the control group with a frequency of 3 compared to 1.1 in the treatment (*Table 6.2*). This is supported by the summary of pupils' responses from the treatment group to question 1 (*Table 6.1* Appendix 6F) which suggests that the change to teaching was noticeable to the point that it outweighed the differences in subject content.

Teaching strategies

Both groups of pupils mentioned a number of teaching strategies, the list was longer for the control group with 13 strategies mentioned compared to 5 in the treatment (See Appendix 6C). This is supported by the data in *Table 6.2* where of the teaching strategies mentioned by the treatment group 39% of them (15) were group work. In contrast no strategy received more than 3 mentions in the control group (See Appendix 6C).

6.5 Section B - Analysis of teacher comments made after teaching an inquiry-based scheme of work

Teachers who had taught using the inquiry-based scheme of work were invited to take part in the focus group. Four teachers (80%) agreed to take part. They were asked a series of questions as detailed in *Table 6.3* and an audio recording was transcribed (See Appendix 6D)

6.6 Analysis

Table 6.3 (see Appendix 6G) summarises the groups' responses to each question separated into comments given within a positive or negative context.

6.7 Findings for teacher perceptions

Attainment and critical thinking skills

Table 6.3 shows that all teachers agree that the use of inquiry-based teaching strategies should be incorporated into teaching practice as they believe that they have a neutral effect on learning subject knowledge and 75% believed that they do have a positive effect on critical thinking skills.

Group work

50% of teachers felt that group work was a the drawback of this scheme of work as it allowed some pupils not to engage, however 50% made comment that this could be overcome with training; it was also suggested that pupils may coast in normal teaching practice (See Appendix D).

Teachers' beliefs about pupil perception

Group work was also given for the reason why teachers (75%) believed that some pupils did not enjoy this style of teaching, the other being given as the absence of a summary of subject facts.

6.8 Section C - Analysis of the themes arising from a teacher focus group compared to similar themes discussed by pupils

A thematic comparison was then carried out against treatment pupil responses. The categories were decided by firstly colour coding the transcript (See Appendix 6D) to identify the themes and then a semi-quantitative analysis with supporting quotes was produced. Themes from the pupil transcript analysis (see Appendix 6C) were then placed beside similar teacher themes for comparative analysis to take place (See Appendix 6E). A summary of this data compared to the pupil data is shown in *Table 6.4.*

Table 6.4 Quantitative Comparison of the Themes Discussed by the Teachers and Treatment Pupil Focus Group

Theme	Teacher group (4 teachers)			Pupil treatment group (7 pupils)		
	Positive *(Per teacher)*	Negative *(Per teacher)*	Neutral *(per teacher)*	Positive *(Per pupil)*	Negative *(Per pupil)*	Neutral *(Per pupil)*
Teacher perception	19 *(4.75)*	18 *(4.5)*	0 *(0)*	0 *(0)*	0 *(0)*	0 *(0)*
Pupil perception	12 *(3)*	10 *(2.5)*	0 *(0)*	25 *(3.6)*	16 *(2.3)*	0 *(0)*
Teaching strategies	13 *(3.25)*	14 *(3.5)*	0 *(0)*	15 *(2.1)*	19 *(2.7)*	1 *(0.1)*
▫ Group work	0 *(0)*	5 *(1.25)*	0 *(0)*	3 *(0.4)*	11*(1.6)*	1 *(0.1)*
▫ Practical	0 *(0)*	0 *(0)*	0 *(0)*	7 *(1)*	4 *(0.6)*	0 *(0)*
Attainment /completion of work	0 *(0)*	0 *(0)*	9 *(2.3)*	4 *(0.6)*	1 *(0.1)*	0 *(0)*
Critical thinking skills	4 *(1)*	0 *(0)*	4 *(1)*	0 *(0)*	0 *(0)*	0 *(0)*

6.9 Section C Analysis and Findings for pupil and teacher perception comparisons

Inquiry-based teaching style

The balance in the teachers' perceptions of this style of teaching is shown in *Table 6.4* by the almost equal number of positive and negative comments. This is also true of teachers' beliefs about pupil perceptions where they gave a ratio of 1:0.8 positive to negative comments which was broadly in agreement with the pupil ratio of 1:0.6 albeit the pupils were slightly more positive. This may have been due to pupils' positive perceptions of practical work and topic content which contributes 48% of their positive comments, these two areas were not commented on by teachers so may not have factored into their conjecture.

Teaching strategies

This balance of positive and negative comments continues as teachers frequently mentioned the idea that a mix or variety of teaching strategies would be optimal. This is indicated by the equivalence of the frequency of their comments in this section, producing a ratio of 1:1.1 positive to negative. Pupils were slightly more negative when commenting on teaching strategies with a ratio of 1:1.3 positive to negative, however of the negative pupil comment frequency of 2.7, 1.6 were contributed to by comments about group work whereas the figure for teachers was 1.25. This indicates that teachers underestimated pupils' negativity toward group work.

Attainment and critical thinking skills

Direct comparisons for attainment and critical thinking skills are more difficult. Teachers were either positive or neutral about the impact of inquiry-based teaching on critical thinking whereas pupils did not mention this. Teachers believed that this style of teaching had a neutral effect compared to other styles on attainment; this is largely supported by pupils where a 4:1 ratio of positive responses indicated that they still completed all of their work, although these comments did not directly relate to their attainment.

6.10 Discussion

Discussion on group work

Pupils' perceptions of inquiry-based teaching strategies when compared to the control can be distilled down to the fact that they did not enjoy group work, as shown by the number of negative comments, highlighted in this example "'Cause we get put into groups where you have very low level science and very high level science, and the lower level expect all the higher level to do all the work…" (See Appendix B, p. 4). This was an agreed theme albeit with slight differences in their reasoning (See Appendix B) and is supported by Wellington (1998) who writes that "It has been claimed… when group work is closely observed and analysed it often reveals domination by forceful members, competition, lack of engagement from some…" (p. 8). Teachers also felt that the increased use of group work encouraged some pupils to coast although other teachers suggested that coasting takes place in other forms of teaching, and that for the strategy to be successful the pupils need training.

As outlined in section 1.4, ability is a term used to define a crude measure of a pupils' current point of learning and therefore expected final GCSE grade. As discussed, it does not take account of the myriad of complex factors that may affect whether an individual will meet the expected level of progress over five years. One very important determining factor is a pupil's disposition towards learning, which will influence their attitude towards learning which in turn may manifest itself in their completion of tasks, as long as these are differentiated to their ability level. When other pupils comment on lower ability pupils they are doing so from two pieces of evidence: one is that they may know the ability band of the pupil, as this is not hidden due to differentiated tasks and levelled assessments; the other is their experience of being in a class with that pupil and analysing their attainment scores. In many cases they are broadly correct in estimating the ‘ability’ of the pupil, and (as in this case) may link it to the pupil’s motivation to complete or assist in the completion of tasks. Although this assumption may well not be correct, it is, in the researcher’s experience not an uncommon pupil perception.

This belief is supported in literature with Crawford, Krajcik, & Marx (1999) asserting that in a middle school class “…it took 8 weeks to become fully productive in working together toward a group goal” (p. 720). This was commented on in the SCORE report (2008) which suggests that teachers should “..pay more explicit attention to the encouragement of interpersonal skills through working in groups.” (p. 8). One pupil did suggest that it was the mixed ability nature of the group that was the issue; “I think next year when we get put into sets it will be better, people will be more dedicated to work…” (See Appendix B, p. 4). However, she did make reference to this having a negative effect on lower ability groups, which is of interest as White and Frederiksen (1998) found that after pupils engaged in a programme designed to improve the metacognitive process in the inquiry cycle, higher ability pupils made more gains in their understanding of physics concepts, whereas lower ability pupils made comparatively greater gains in their inquiry skills suggesting the strategies are effective across the ability range.

Discussion on critical thinking skills

The teachers’ assertion that inquiry-based strategies have a positive effect on critical thinking skills is in line with literature that discusses teachers’ beliefs (Wallace & Kang, 2004; Wilson and Mant, 2011; Minner, Levy, & Century, 2010). Wallace and Kang summarise that “Teachers intuitively believe that inquiry-based activities will promote scientific thinking skills, but as yet they have little evidence to support this hypothesis.” (p. 959). This was true of the findings of this research, as despite 100% of the teachers believing that the value of inquiry-based strategies was on critical thinking gains, there was found to be no statistically significant increase in critical thinking skills in either group during the experimental period in chapter 4 (*Table 4.3 – 4.6*).

Discussion on pupil training in inquiry-based learning

It is also asserted that teachers who have experience of or favour this style are much more likely to make effective use of it and therefore report gains (Brown & Melear, 2006; Geier *et al.,* 2008; Wallace & Kang, 2004; Crawford, 2000). This was true of the teachers interviewed for this research where two teachers were far more positive about its use. The skills mentioned by one teacher showed that he had a good understanding of what critical thinking skills are as they reflect the list in Chapter 2, he went on to suggest how pupils could be better trained to benefit from it;

"It's that thing again, because of lack of training, it's a good thing to do but it does take a lot of time for the kids to get this idea of; they're responsible for the learning, they're responsible for taking out the information and assessing it, deciding what's important and what's not." (See Appendix D, p. 3).

The importance of pupil training also appears in literature (Crawford *et al*. 1999; Keys & Bryan, 2001). Frederiksen and White (1998) used a computer programme to teach pupils the stages of scientific inquiry explicitly before pupils applied these ideas, as they found that pupils could not see the purpose of each stage until they had taken part and reflected on it. This led to the conclusion that inquiry-based learning needs to be implemented early and regularly. Hmelo-Silver, Duncan, and Chinn (2007) suggested that scaffolding of skills could be applied to assist in pupil training in their response to Kirschner, Sweller, and Clark (2006) who assert that the use of constructivist approaches can lead to pupil misconceptions and frustrations, a position also supported by Kuhn, Black, Keselman, & Kaplan (2000) and Klahr and Nigram (2004). These ideas are supported by pupil comments as attainment is discussed.

This notion of training pupils or assisting them in learning the inquiry skills is supported by both Furtak *et al.* (2012) and Lazonder and Harmsen (2016). In both these meta-analyses higher levels of guidance positively affect pupil outcomes. In the case of Lazonder and Harmsen this is seen in "performance success" (p.11) as pupils achieve more and can evidence work in lessons. It is also found that although all types of guidance can improve pupils' learning outcomes the effect size is not significant, and thus it is the inquiry skills that are better developed through guidance rather than the knowledge base. In the analysis by Furtak *et al.,* again it is the procedural, epistemic and social categories of cognition (p. 322) rather than the knowledge base ("conceptual framework") that are better served by inquiry-based learning, which is in turn positively affected by a more teacher-led approach. However, as previously stated in section 5.8, these reported improvements in inquiry skills do not necessarily lead on to gains in attainment, as was implied by teacher comments. It is also true that although teachers acknowledged the benefits of this type of learning in terms of inquiry skills, pupils did not recognise or mention this area of development.

The study by McConney *et al.* (2014) acknowledges that the PISA tests do not capture all areas of pupil learning that might benefit from inquiry and so suggest that more work could be done into "[identifying] those aspects of inquiry that best promote science learning while positively engaging students" (p. 29). This is in line with the views of the teacher that there were some aspects that pupils enjoyed and benefitted from more than others and perhaps these could be integrated with more traditional teaching methods.

Discussion on attainment

This theme of subject knowledge and understanding was commented on by both teachers and pupils. One pupil stated that "…(I) feel that it wasn't as good for my learning because I've found that I've been struggling to get things to work efficiently so there hasn't been time to do what I like to do to memorise it and to make my own links and knowledge. I haven't had time to do that…" (See Appendix B, p. 5). The idea that pupils like to have a body of facts written down to refer to in order to assist in their understanding and future attainment was also highlighted by teachers with the suggestion that this could be heavily incorporated into any further use of these strategies. The main reason for this being the preparation of pupils for assessments, this teacher perception was found to be held widely in literature (Keys & Bryan, 2001; Wallace & Kang 2004; Mant *et al*., 2007) as a main reason why teachers do not engage in inquiry-based teaching. Despite this, teachers still reported a neutral effect on pupils' attainment from teaching using inquiry-based strategies. This was supported by the results in Chapter 5 which found that in both the control and treatment group pupils did improve their attainment over the experimental period however the data unavailable to teachers was that the comparative attainment change for the treatment group was statistically significantly smaller than the control group and although attainment can be affected by many variables (discussed in chapter 5) the teachers assertion in the case was incorrect; the treatment had a slightly negative affect.

6.11 Drawbacks of method

The formation of the focus groups (discussed section 3.12) meant that it was not possible to get a stratified sample in terms of the teaching groups and the ability range which would make conclusions about the cohort pupils' perceptions more valid. Nine pupils took part and all of them pupils had a National Curriculum Key Stage 3 target of Level 7a or 7b and so represent the top third of the cohort in terms of ability. Therefore, the information gathered and themes discussed only represent this ability range. The semi-structured interview format was used in order to elicit pupils' general thoughts and encourage a wider incorporation of themes; the drawback was that the views of some had an influence on those around them and therefore some topics discussed tended to dominate. However, all pupils were asked each question to ensure that they all had the opportunity to respond. The questions were open and large in scale so that they could be applied to both the control and treatment groups however a divergence in the topics discussed by the two groups did emerge making comparison slightly more difficult.

In the teacher group their understanding of the aims of the research and their involvement in teaching meant that they gave full and detailed answers which fulfilled the aims of the research question. Although the drawback of this was that their involvement in my research from the start may have encouraged demand characteristics to affect their responses due to their relationship as colleagues. Furthermore, social acceptance may have played a part as when commenting on teaching with other professionals (including the head of department) teachers may have felt that their responses would reveal information about their teaching behaviours and encourage judgement.

6.12 Discussion on Major Themes

This discussion suggests that teachers identify advantages of inquiry strategies more so than pupils who did not mention any other skills (other than subject understanding) that they may be developing. This could be because these skills are not highlighted as being a product of the UK educational system and they are never tested, so the importance of their development is diminished. Practical work was seen as a positive aspect of teaching in both pupils groups, but comments on other aspects of investigative work e.g. planning were rare. Pupils value the collection of a body of knowledge possibly due to assessment demands and although teachers see the value in developing critical thinking skills, they too indicated that summarising key facts was important. It is also indicated that the social and learning skills required to work collaboratively and independently on investigative work are underdeveloped in the pupils involved in this research, which leads to frustration in pupils who feel that they want to learn and are held back by others. Teacher believe that the integration of inquiry-strategies into traditional direct-instruction forms of teaching is beneficial but training needs to be in place to support pupils to develop the skills necessary for these to be effective.

Chapter 7- Conclusions

7.1 Introduction

In the following chapter the findings of this research will be outlined and then discussed with reference to current literature. The limitations of this research will also be considered in regard to the validity of conclusions that can be drawn, and the effect of confounding variables on these conclusions. Finally improvements to the research practice and suggestions for further areas of research will be discussed.

7.2 The Research Questions

This was a mixed methods research project as detailed below:

Research Question 1 was 'What is the impact of an inquiry-based scheme of work on pupils' critical thinking?' and was measured using a quasi-experimental intervention design, with a pre- and post-test of critical thinking for pupils in intervention and treatment groups. It was found that although both groups' critical thinking skills increased, this increase was not statistically significant. However, the power analysis suggested that, with an effect size of 0.212, significance could have been found with a cohort double the size for the treatment class taught by the researcher. The pre- to post-test score change, although larger in the treatment group, was not statistically significantly larger than the control group. This suggests that there is no evidence to support the hypothesis that teaching using inquiry-based strategies has an impact on pupils' critical thinking skills.

Research Sub-question 1a was 'Is there a delivery variable that influences these effects?' and was answered by comparing the data of pupils in researcher's class to the treatment group as a whole and the control group. It was found that because of the statistically insignificant outcomes discussed above that no evidence was found to support the hypothesis that a delivery variable, of the researcher as the teacher having, affected the findings of this question.

Research Question 2 was 'What is the impact of an inquiry-based scheme of work on pupils' attainment?' and was also measured using a quasi-experimental design, with a pre- and post-test for pupils' attainment in control and treatment groups. It was found that an increase in attainment scores found for both groups was statistically significant. It was also found that the increase in attainment score from the pre- to post-test was statistically significantly greater in the control group. This evidence does not support the hypothesis that teaching using inquiry-based strategies has a positive effect on pupils' attainment.

Research sub-question 2a was the same as 1a and measured in the same way. In contrast to the findings in the last paragraph, it was found that the control group did not have a statistically significant increase in score greater than the researcher's treatment class. Therefore this provides evidence that having the researcher teach one of the treatment classes could be a delivery variable which affected the findings of this research question.

Research question 3 was 'What are pupils' and teachers' perceptions of the use of inquiry-based teaching strategies in science?' and was investigated using three focus groups, one pupil from each of the control and treatment groups and one teacher group. These were run using the semi-structured interview format. It was found that pupils in the treatment group did not enjoy mixed ability group work but pupils in both the treatment and control groups did enjoy practical work. Teachers agreed that they believed that inquiry-based teaching does improved critical thinking skills but that this teaching style would have no impact on attainment, as measured currently. Pupils said they struggled with the independence of investigative work as part of a group and teachers suggested that pupils needed more training in these skills.

7.3 Addressing the Literature

Gaps being addressed

This research was carried out as existing research suggests that practitioners believe that teaching using inquiry-based strategies is of benefit to pupils (Wallace & Kang, 2004; Minner, Levy, & Century, 2010; Wilson & Mant, 2011). This was supported by the teachers in the focus group who all agreed that these strategies should be incorporated into the current curriculum to help improve pupils' thinking skills.

A further justification was that education policy explicitly states that pupils should be taught the 'working scientifically' strand of the curriculum (DfE, 2013) which includes: "Ask questions and develop a line of enquiry based on observations of the real world, alongside prior knowledge and experience" (p. 4). Therefore, teachers must provide the opportunity to develop an inquiry-based scheme of work which enables teachers to meet this requirement. Support for these methods of teaching can be found in other reports that have been funded or presented to the government (NESTA, 2005; ASE, 2009; Science select committee, 2006).

A further gap in the literature was a lack of agreed definition as to what inquiry-based learning was, and whether the skills outlined in the National Curriculum could be taught by direct instruction rather through their use (Yeomans, 2011). Finding an agreed definition relied upon research that was already in place internationally and in the UK alongside the reports detailed above. This definition provided the opportunity to meet National Curriculum standards, thereby providing a valid context and justification, in a way that was not already in place in the research school. As research in the UK is limited, compared to that carried out in the US (White & Frederiksen, 1998; Crawford, Krajcik, & Marx, 1999; Crawford, 2000; Brown & Melear, 2006; Hmelo-Silver,Duncan, & Chinn, 2007; Wallace & Kang; 2004; Minner *et al*., 2010), the definition arrived at was informed by these US papers alongside a derivation of the skills outlined for development, in the reports listed above alongside the National Curriculum (2013). With this definition it was then possible to develop a scheme of work to meet both the subject content and the working scientifically strands.

7.4 How this research addresses the literature

The findings of this research were that teaching exclusively using inquiry-based strategies had a negative impact on attainment scores from a test which incorporates questions on the ‘working scientifically’ strand. This may indicate that Yeomans (2011) and Ryder (2011) are correct in their assertions that learning through inquiry is not the same as learning inquiry skills, a position that was supported by the findings of Klahr and Nigram (2004), and therefore, perhaps the attainment of the control group in this research was achieved as pupils were taught the inquiry skills that were likely to be assessed rather than, as in the treatment group, learning through use of them. It is also true that the test was assessing pupils’ subject knowledge so a lower attainment may be due to inquiry-based strategies being less effective at preparing pupils for the traditional tests that are used in school. This conclusion is supported by Horner (2011) when stating “Methods of assessment do not…reflect the integration of knowledge and inquiry.” (p. 14) and therefore teachers believe that taking time to learn by these strategies removes time for exam preparation (Keys and Bryan, 2001; Wallace & Kang, 2004). The findings of research question 3 in this research also support this, as all teachers interviewed did not think that inquiry-based learning would have a positive effect on the attainment of pupils. They made suggestions that pupils missed having a body of facts that they could learn for a test (discussed in section 6.22). What this suggests is that current teaching suits these tests, as it has evolved with an understanding of how pupils are going to be assessed.

This is also supported by the meta-analyses by Furtak *et al.* (2012) and Lazonder and Harmsen (2016) as neither found that inquiry-based learning had a positive effect on measures closest aligned to the attainment measure in this study, “conceptual structures” in the former and “learning outcomes” in the latter. This was true when differences in both degrees of teacher guidance and pupil age were taken into consideration.

However, this does not mean that learning through inquiry-based strategies is not beneficial to the education of pupils, as shown by the findings of Research Question 2 which indicate that it was not of benefit. Other literature reviewed has found a positive impact of inquiry-based strategies on attainment using similar standardised tests. (Geier, Blumenfeld, Marx, Krajcik, Fishman, Soloway, Clay-Chambers, 2008; Mant, Wilson, & Coates, 2007) which was also true of the treatment class in this research that was taught by the researcher, presented in Sub- question 2a. In line with the research carried out by Geier *et al.* and Mant *et al*. a suggestion for further research and/or implementation would be that teachers need more training and involvement in the development of the resources. This could help inquiry-based strategies to be as effective at raising attainment as the 'traditional' style of teaching which was demonstrated by the control group in this research. This suggestion is supported by conclusions made by others (Yeomans, 2011; Wallace & Kang, 2004; Geier *et al.,* 2008; Brown & Melear, 2006), who suggest that teachers need to understand the pedagogy of this type of teaching, be assisted to understand and develop resources, and that finally, they must have a belief that pupils are capable of learning through these methods.

Despite much literature stating that pupils like being taught by both inquiry-based strategies and group work (Murray & Reiss, 2005; SCORE, 2008; Wilson & Mant, 2011) the findings from research question 3 did not support this. Pupils struggled to work independently in mixed ability groups and felt that this had been of detriment to their learning (discussed in section 6.22). Although pupils did make some positive comments about discussion in groups, the pupils interviewed (all higher ability) found the strain of helping weaker students detracted from this enjoyment. It was also the case that they were positive about practical elements, as was also found in the research listed above, however they preferred previous topics where they were given a practical to do rather than, as part of this research, having to plan the investigations in groups.

This supports the work of Crawford *et al.* (1999) who found that it "took up to 8 weeks to become fully productive in working together toward a group goal." (p. 720) and is further developed by Wellington (1998) who asserts that within group work there is "domination by forceful members [and a] lack of engagement for some" (p. 8). Therefore as suggested by the teachers in the focus group reported for research question 3, pupils need training in how to work as a group and apply these skills. If these skills are developed early then all pupils should be able to contribute and the pressure on pupils' use under-developed inquiry skills, alongside social skills and the development of scientific understanding would be much reduced. This is in line with the response of Hmelo-Silver *et al*. (2007) to Kirschner *et al.* (2006), where they suggest that although there is an increased cognitive load when learning through inquiry-based strategies, this can be mitigated through scaffolding, sections of direct instruction and well-structured tasks, a position supported by Kuhn and Pease (2008), White and Frederiksen (1998) and Millar (2004).

This is also in agreement with the meta-analysis carried out by Furtak *et al.* (2012) who found that teacher guided inquiry had the greatest positive effect size when compared to traditional teaching (p. 319). They also found that the social interaction element of cognition is improved by inquiry learning (p. 322) and so pupils' ability to work as a group and communicate scientifically should improve over time. In support of this enhanced emphasis on the degree of teacher-guidance, Lazonder and Harmsen (2016) found improvements in performance success (p. 20) across all ages, which may lead to an increase in work completed by all abilities and as such solve the problems the higher ability pupils reported facing in mixed-ability group work.

The findings of research question 1 are that, in this research, inquiry-based teaching strategies have no effect on critical thinking skills. Although there was an improvement in both groups, and a much greater effect size in the treatment group, none of these increases were statistically significant. This was in opposition to the views of teachers collected for research question 3, who thought that this style of teaching would benefit pupils' critical thinking skills, a view shared by other teachers in research (Wallace & Kang, 2004). This does not mean that the study has nothing to add to the body of research in this area. A power analysis showed that for the largest effect size, the treatment class taught by the researcher, a larger cohort could have shown a statistically significant increase suggesting that further larger studies could be carried out. This is supported by research carried out using the same critical thinking tests (Cornell Series X) which found positive gains with larger sample sizes (Ernst & Monroe, 2004; Daud & Husin, 2004). It is also a common theme in research that it takes time to develop critical thinking skills (Miri, David, & Uri, 2007; Halpern, 2001; Marin & Halpen, 2011) and therefore, the findings from this research may have shown more positive gains if carried out over a longer period of time.

The meta-analyses by Furtak *et al*. (2012) and Lazonder and Harmsen (2016) suggest that benefits of inquiry-based learning can be found in skill areas that could be aligned with the critical thinking skills measured in this research. Improvements were found in the cognitive areas that control developing and evaluating evidence, the communicating of science and the ability to ask questions and design experiments (Furtak *et al.* p305), whilst in Lazonder and Harmsen (2016) these improvements were seen in the ability to complete tasks including using inquiry skills (p. 20).

One positive element that this research can add is that it is possible to measure the critical thinking skills of lower secondary age pupils, and that the test used provided both practicable and useful measurements that could be compared over a period of time. The promotion of critical thinking skills has been one encouraged by the UK Department for Education (DCSF, 2008) and therefore further use of this test may raise the importance of teaching these skills amongst teachers, as research suggests that (Ku, 2009) teachers focus mainly on the elements of knowledge and skills that will be assessed.

7.5 Limitations of Research

Findings

As the summary of findings from research question 1 indicates there is no evidence of the impact of inquiry-based learning on critical thinking skills. However, as effect sizes were found, it is possible that with a larger sample size a different conclusion would be reached. It is also possible that any impact may have been limited by time, and therefore, the size of the gains that were measured.

Conclusions on the negative impact of teaching by inquiry-based strategies on attainment, from research question 2, as with all the conclusions from this research, are only applicable to the research school context and it is also worth noting that the negative impact was only found in two of the three treatment classes. This may be as the third class was taught by the researcher or the quality of teaching provided by the other two teachers. It is also true that there was much variation in pupils' attainment increases within the two groups, suggesting a number of variables e.g. teacher commitment, pupil control, familiarity of teacher with the teaching style, teacher expertise in the subject matter and question scaffolding were affecting the research. Therefore, some caution should be applied when extrapolating these conclusions to a different context.

The delivery variable, discussed in research question 1a and 2a, of the researcher being one of the treatment groups' teachers suggests that this may have been a contributor to the positive gains that this class showed in both attainment and critical thinking skills. Although, one teacher is a very small sample it never the less offers a clear suggestion for further investigation into how teachers trained and invested into teaching using inquiry-based strategies may impact pupils in these two measures.

The conclusions from pupil focus groups in research question 3 were particularly skewed as, despite the researcher's efforts, all the opt-in volunteers were in the higher ability band. Therefore, all findings on their perceptions of inquiry-based teaching strategies are limited to this group. The shared perception of not liking group work due to helping or doing the work for lower ability pupils suggest that the findings would not be generalizable to the entire cohort within this research, and therefore, only relatable to higher ability pupils with this school context.

7.6 Confounding Variables

The attainment tests used were of the style that is routinely used in the research school. As such, teaching had adapted over time with knowledge of the types of questions that will appear. The teachers of the control group were teaching as they would normally teach, therefore, it is highly likely that types of questions will have been referred to or even that parts may have been used as lesson activities. This means that the control group would have had an advantage when completing the post-test, leading to an expectation of a higher increase in attainment over the treatment group.

The critical thinking tests used were suitable for the age range of the cohort; however the research cohort were at the bottom of the recommended range. This suggests that pupils who had a reading age below that of a standard Year 8 pupil may have struggled with the paper. This led to far fewer papers being completed in the entirety in either the pre- or post-test for lower ability pupils, and may have skewed the data upwards. It was also the case that some classes had a much lower completion rate than others indicating that teacher encouragement of pupils was a variable that could not have been controlled but will have affected the data causing an uneven distribution of completed tests in the control and treatment groups and a much smaller sample size for analysis.

Teachers' beliefs and knowledge of the inquiry-based pedagogy was a variable that will have caused variation in classroom practice, and therefore a major confounding variable was the fidelity of delivery of the intervention. As some teachers liked and saw the benefit of this style of teaching more than others, it may have led to a greater motivation to reveal its impact. In order for the lessons to be truly inquiry-based it was not just the use of resources and following of the lesson plan that was important, it was the pupil interaction, providing scaffolding and expert questioning which required teachers to be considered and precise in their interaction with pupils. As this area of teaching could not be controlled, and taking into account the varied perceptions and expertise of the teachers involved, it must be accepted that this will have caused both variation within the treatment group and a possible reduction in positive impact on the two quantitative measures.

7.7 Implications and Opportunities for Further Research

The literature that has been quoted throughout this concluding chapter shows that a similar investigation could be carried out but with a number of changes. The study could be over a longer period of time in order for larger changes to be seen and critical thinking skills to develop. The cohort should be larger as effect was shown in this research but it was small and therefore a larger sample could test whether these effects were real; this could be carried out either across year groups in the same school or similar age groups in different schools to improve quality of data and make findings more generalizable.

As the research suggested that the teacher delivering the intervention had an effect on outcomes, a larger sample size would provide the opportunity to invite teachers who are interested in teaching using inquiry-based strategies to be involved in the research. Teachers could be trained in this pedagogy and its application in the classroom. Resources and lesson plans could be developed collaboratively so that all treatment classes are receiving a more similar experience (compared to this research), thereby improving the fidelity, as teachers are more motivated towards positive outcomes. The two test measures could also be developed as detailed below.

Attainment tests are in a state of flux in lower secondary school science at present in England. The government has terminated the use of the National Curriculum levels as from September 2016. This means that schools have a lot more autonomy on how best to measure the progress of pupils from key stage 2 to 4. This progress still has to be measured and reported in order to hold teachers and schools accountable for the outcomes achieved by pupils throughout this period. This autonomy provides an opportunity to uncouple the learning of science from the traditional test. Attainment can be measured in different ways, incorporating skills and deeper levels of understanding, giving opportunities for research into subject knowledge without the constraints of the attainment tests used within the school context.

The Cornell Critical Thinking Test (CCTT) was one of two tests of critical thinking that have been peer-reviewed, have precedent in literature and were written for the age range of the cohort. However, further research into this area could lead to the development of a test which places lower secondary science pupils in the middle of its age range so that all pupil abilities could access the questions equally. It would then also be possible to trial a range of tests to see which was the most suitable for the cohort and produced results with a valid distribution.

One of the biggest limitations was pupils' ability to work effectively in groups; this was suggested by both the teachers and the pupils in the focus groups. Therefore, further research could be carried out into literature surrounding pupil group work; leading to an implementation of strategies to develop these skills in pupils embarking on a solely inquiry-based curriculum. The other drawback for pupils was accessing the skills needed for inquiry; therefore, research could be carried out into whether actively incorporating the direct teaching of inquiry skills before or alongside an inquiry-based programme would improve the outcomes in attainment and critical thinking skills from these pupils.

Appendix 3A

This table differentiates teaching activities that may and will not be incorporated into an inquiry-based teaching strategy.

Table 3.1

Inquiry based teaching strategies in Science	**Non-Inquiry-based teaching strategies in Science**
Planning investigations using authentic questions	Planning investigations with a pre-determined ‘correct’ method, equipment and outcome.
Carrying out practical work to discover phenomena based on prior understanding of that phenomena and concluding on results actually achieved.	Following a ‘recipe-style’ method to reproduce an expected phenomenon, relying on inductive methods to generate understanding.
Group discussion work in which practical work or secondary data is used to construct understanding.	Group work in which pupils are not cognitively active and are provided with the answers.
Pupils are encouraged to find the answers to theirs or their peer’s questions and provided with the resources to do so.	Pupils are told the answer by the teacher immediately.
Teacher acts as a facilitator or information source at appropriate points.	Teacher talks in lecture style from the front.
Pupils generate their own notes from different sources of primary and secondary data, including the teacher and peers.	Pupils copy notes from board, text books or dictation.
Pupils are taught skills that are appropriate to inquiry, such as collaboration, communication, evaluation in an active context.	Pupils are expected to use communication, collaboration and evaluation without a framework for their development.
Pupils peer assess others work and provide feedback based on their understanding gained from collaboration with peers and collectively agreed criteria.	Pupils are teacher assessed and given feedback based on their ability to get information right or wrong, or peer assessed using teacher generated criteria
Topics are taught with an over-arching big question, which can be broken up into smaller ones with an emphasis on skills, and processes.	Lessons are compartmentalised into learning aims focussing on fact building.
Pupils are encouraged to take ownership of their learning. Identifying areas of weakness and working to close the gaps in understanding.	Pupils are told what they need to know and are corrected when mistakes are identified by the teacher.
Pupils progress is measured by the ability to process information and find a logical route towards the discovery of an answer.	Pupils progress is measured by the ability to recall facts.

Appendix 3B

Table 3.2 An Example Of A Critical Thinking Skills list – Reproduced and Simplified

Halpern, 2001, p272	**Adapted for KS3 Pupils**
Reading with a high level of comprehension	Reading with a high level of comprehension
Providing support for a conclusion	Giving examples to support a conclusion
Understanding principles of likelihood and uncertainty	Suggesting errors and identifying risks
Using analogies	Developing visual and physical models and simple analogies
Reasoning about ratios	Understanding and applying simple ratios or percentage chance
Recognizing the difference between correlation and cause	Understanding with examples that correlation does not always mean causation
Combinatorial reasoning	?
Isolating and controlling variables	Understanding that variables which are not the independent or dependant variable must be controlled
Evaluating evidence	Simple two sided evaluations of argument
Planning a course of action	Planning a course of action
Generating hypothesis	Writing predictions
Using retention (memory) strategies	Developing and using retention (memory) strategies
Making spatial representations	Making simple spatial representations
Restructuring problems	Suggesting different ways of dealing with a problem
Using problem-solving heuristics	Using common sense and experiences to help solve problems
Seeking patterns	Identifying simple patterns
Incorporating anomalous data into a coherent framework	Identifying anomalous data, suggesting reasons for it and removing it.
Recognizing regression to the mean	Carrying out repeats and recognising those which are close to the mean

Appendix 3C

Microbes Inquiry-Based Scheme Of Work

	Lesson Outcomes	Lesson Order	Notes	Resources	Science Process
1	• State what a micro-organism is • Describe some negative factors associated with living alongside micro-organisms • Describe the conditions that micro-organisms need to survive • Design an experiment using Glo-germ to explain the impact of variation in washing hands method	**Authentic big question Why is it important to wash your hands regularly?** Take some initial suggestions but don't pass any information – move them quickly on to the next task. Collect ideas from the group about what a micro-organism is and the conditions that it needs to survive from each other and information provided – This should be on a large piece of A3 people – tidiness is not important, ideas are. Generate a class discussion on their answers and then play the clips about epidemiology. Get them to discuss in groups why this is important to answering the big question. The information sheet has been pre-dusted with glitter—bug. Turn the lights off and use the UV light over their hands to show them how easy it is to spread bacteria. Emphasise that glitterbug is a MODEL for bacteria. Introduce the idea of the investigation plan. Get pupils to plan their investigation, wandering around and supporting their ideas without thinking that there is a particular way that you want them to do it. Problems will be discussed in the evaluation.	Nominate or get them to vote on a leader – this person is also responsible for the on task behaviour of their group. Credits accordingly. Encourage collaboration and stronger readers to read information out to the group. Glitterbug is expensive and only a little amount is needed so please be careful. This should be on a large A3 piece of paper – this needs photocopying onto A4 for next lesson so each group member has a copy in their book. A3 planning sheet – needs photocopying onto A4 for next lesson. Equipment order needed from pupils (you could predict most of this)	Power point – Projector with speakers Emma's Microbe scheme of work – lesson 1 A3 paper, felt tips, Laminated information card pre-dusted with 'Glitterbug' powder. UV lamp Plan frameworks – Large A3	Collaboration Science careers Research Planning
2	• Carry out an investigation,	**Authentic big question Why is it important to wash your hands regularly?**	Apart from obvious unsafe or wasteful practice – allow the pupils to carry out	A4 photocopies of mind	Practical skills Collaboration

	collecting data and controlling variables • Describe the results of the investigation • Explain how the results could be related to a real-life situation when dealing with micro-organisms • Evaluate the method of the investigation	Pupils receive back their own copies of science content mind map and planning sheet from last lesson and stick into book. Five minutes to discuss and distribute tasks in group before carrying out investigation. Collect data in books and then complete the end of the planning sheet in groups. Generate a discussion that evaluates their experiment – try to get others in the class to make suggestions on how they could have improved it. Link back to subject knowledge objectives from last lesson – Get pupils to assess themselves on the outcomes and write a target for next lesson – subject knowledge from lesson 1 or science processes from lesson 2	their investigations. Well- timed questioning and support will be useful this lesson.	map and planning sheet for each pupil. Equipment pupils ordered Glitterbug is expensive and only a little amount is needed so please be careful.	Data collection Explaining results Evaluating
3	• State what unicellular means • State three types of micro-organisms categorised by their method of nutrition • Investigate and describe how saprobionts, symbionts and photosynthesising micro-organisms gain nutrition	**Authentic Big question: How are micro-organisms adapted to gain nutrients?** Pupils need to move into their groups and answer the questions on slide 2 of the powerpoint. They will need scrap paper and pens. They will need a leader. Emphasise in that at each station – one person needs to read the task out loud to their group. (those with literacy issues will need supporting) – **Station A will be at the class computer.** Generate a class discussion on these questions, acting as a facilitator and drawing on the knowledge in the room. Introduce the actual question that they are	Emphasise their knowledge and act as a facilitator, prompter and co-collaborator. The stations will arrive in trays so need spacing around the room. You will need to explain the use of a microscope, health and safety and tidying up at each station. Give literacy support to those who require it. Take booklets in at end of the lesson – as some are likely to lose.	Workbooks will be photocopied – ask Heather for them. Order Micro-organisms – Emma's scheme Lesson 3 and 4	Practical skills Collaboration Data collection Explaining results

	• Explain how micro-organisms are adapted to gain nutrients	going to investigate and hand out the 'Big question' work books – Talk them though the sections of these. Emphasise the group work and discussion – You can give positive credits for these interactions. – **this could take up to 20 mins** They need 10mins –ish for each station (over two lessons) but judge the room. They need to work in groups and fill in their booklets as they go.			
4	• State what unicellular means • State three types of micro-organisms categorised by their method of nutrition • Investigate and describe how saprobionts, symbionts and photosynthesising micro-organisms gain nutrition • Explain how micro-organisms are adapted to gain nutrients	**Authentic Big question: How are micro-organisms adapted to gain nutrients?** Continue on from last lesson. Encourage workbooks to be filled in as they go along. When finished they can do the organisation of their ideas pages and the concept map. This could be completed for homework. **Last 7 mins** - Refer them back to the learning outcomes at the end. Get them to discuss with their groups their understanding of each one and then write a target for themselves in relation to the subject matter.	Stick booklets into exercise books.	Same as above.	Same as above.
5	• State the chemical that is required for living organisms to respire and release energy	**Authentic Big Question: What affects the growth of micro-organisms?** In groups - Discuss and research the two chemicals needed to release energy, and the name of the process (they should have covered this in the	Keep hold of plans for next lesson. Scaffold questioning with the groups both of the science processes of planning and the scientific information.	Order Micro-organisms – Emma's scheme Lesson 5 and 6 – this will include	Collaboration Science careers Use of microscopes Research Planning

	• Describe the two ways that this chemical molecule can be gained • Explain the chemical reaction which all living organisms use to release energy • Plan an investigation into the effect of a factor on the growth of micro-organisms	respiration topic). Also what the micro-organisms need to release energy for. In groups – Look at the budding yeast down a microscope – ask them to draw on whiteboards what they see. Introduce the Big question. Generate a class discussion and help them make the link between energy and growth through reproduction. Explain that they will be choosing a question that will help them to investigate the big question. Provide them with an information pack (ISA-style) in groups and ask them to make a group plan. Plenary – get groups to pin up their plans and get the other groups to move around clockwise read and write improvements on post-it notes and stick them on plan.	Lead them towards a method that will work with quantities but do not tell them. Do not complete final section this lesson. Order the equipment that the pupils need for next lesson. Take in A3 plan and reduce down to A4 making a copy for each member of the group ready for next lesson.	photocopying You need the microscopes	
6	• State the chemical that is required for living organisms to respire and release energy • Describe the two ways that this chemical molecule can be gained • Explain the chemical reaction which all living organisms use to	In groups – on A4 copies of plan pupils individually improve them. They then discuss the method – and assign roles to members of the group. They carry out their investigation. Pupils complete the bottom section of their plan. Refer back to the learning objectives from the previous lesson – Discuss with the pupils what they have found and how it relates to the outcomes.		The equipment that the pupils order in separate trays.	Practical skills Collaboration Data collection Explaining results

	release energy • Plan an investigation into the effect of a factor on the growth of micro-organisms				
7	• State three ways in which micro-organisms are useful to humans • Describe the roles of micro-organisms in digestion, photosynthesis and fermentation • Explain why micro-organisms roles in these processes are essential to humans	**Big question: Why are micro-organisms important to humans?** Pupils start a spider diagram taking a double page in their books – they can discuss their ideas with their groups. After about 5 mins add the stimulus questions one at a time, state them verbally to give them further ideas for additions. Introduce the task, in groups they move around to each station. Emphasise that there will be five minutes of discussion before 5 mins of note taking. Pupils return to their groups. The final questions can be started in class and finished for homework. There is a developing sheet for the very weakest if required.	During the lesson you must act as a facilitator by scaffolding questions and directing pupils to information or other group members. Encourage them to look up words in the textbook and highlight keywords on their spider diagrams. Try to stick to the timings. This homework can be taken in and marked and returned with successes and targets.	Order micro-organisms – Emma's scheme Lesson 7	Collaboration Research Analysing graphs
8	• State some features of a micro-organisms cell that are useful • Compare a micro-organisms cell to an animal or plant cell • Suggest how some adaptations of cells make	Introduce the big question **Why do animal, plants, bacteria and fungi cells look so different and how is this used to benefit humans?** Pupils (in groups) try to draw an animal and a plant cell with labels. After they have had a go put up the keywords to see if they can get any further. Then ask them to leave what they have done and put it to one side until later in the lesson.	Encourage group discussion and motivation- this task should take about 25 mins. Re-emphasise line	Order micro-organisms – Emma's scheme Lesson 8	Collaboration Research Practical work Diagram drawing

	them useful to humans	Introduce learning objective and outcomes. Set them off on task 2 – using bio-viewers and photocopies from textbook and GCSE bitesize. Encourage them to highlight key information and make links between the structure and its role in the cell. Use slide 4 to generate class discussion on the difference between the cells – and the linking of this to the big question. 5 minute discussion in groups as a stretching activity on humans using the features of micro-organisms. Facilitate their thoughts and if time have a class discussion.	drawing in science – no shading – no sketching – large. Try not to give answers – bounce them back to others in the class. Act as a stimulator and co-ordinator. FYI – fungi have a protein cell wall – so good as a protein substitute. Bacteria have DNA that is easy to extract (not in a nucleus) and they reproduce quickly.		

Appendix 4A

Table 4.3 *Descriptive statistics comparing the critical thinking scores of the control and treatment groups, including a test to determine normal distribution*

	Control		Treatment		Treatment class taught by experimenter	
	Pre- test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Mean	24.4	25.4	27.1	28.7	26.7	29.6
Standard Dev	13.4	14.1	12.7	12.2	13.9	13.5
Shapiro – Wilk p value	.760	.950	.660	.360	.890	.180
Normal	Yes	Yes	Yes	Yes	Yes	Yes

Table 4.4 *Comparison of pre- and post-test critical thinking scores (using a paired t-test) or the control and treatment group and the treatment group taught by the researcher and the control group*

Paired Samples Statistics

Control_Treatment			Mean	N	Std. Deviation	Std. Error Mean
Control	Pair 1	Pre_Test Score	24.4	45	13.4	1.99
		Post_Test Score	25.4	45	14.1	2.10
Treatment	Pair 1	Pre_Test Score	27.1	55	12.7	1.71
		Post_Test Score	28.7	55	12.2	1.64
Treatment (Res.)	Pair 2	Pre_Test Score	26.7	30	13.9	2.53
		Post_Test Score	29.6	30	13.5	2.47

Table 4.5 *Correlations for pre- and post-test critical thinking scores (using a paired t-test) for the control and treatment group and the treatment group taught by the researcher and the control group*

Control_Treatment			N	Positive Correlation
Control	Pair 1	Pre_Test Score & Post_ Test Score	45	.756
Treatment	Pair 1	Pre_Test Score & Post_Test Score	55	.748
Treatment (Exp)	Pair 2	Pre_Test Score&Post_Test Score	30	.773

Appendix 4B

Table 4.7 Descriptive statistics comparing the change in critical thinking scores of the control and treatment groups, including a test to determine normal distribution

	Control Change *(Post test score – pre test score)*	Treatment Change *(Post test score – pre test score)*	Treatment Change (Res.) *(Post test score – pre test score)*
Mean	0.99	1.53	2.93
Std. error of mean	1.43	1.19	1.68
Standard Dev	9.60	8.84	9.22
Shapiro – Wilk *p* value	.610	.960	.816
Normal	Yes	Yes	Yes

Table 4.10 Descriptive statistics to ascertain the distribution of pupils' CCTT scores, target and pre-attainment levels

	Target	Pre-attainment	CCTT score
Mean	7.0	5.65	25.9
Standard Dev	0.84	0.68	12.7
Shapiro – Wilk *p* value	.000	.002	.32
Normal	No	No	Yes

Appendix 5A

Table 5.1- *Descriptive statistics comparing the attainment scores of the control and treatment groups, including a test to determine normal distribution*

	Control (n=75)		Treatment (n= 78)		Treatment taught by researcher (n= 27)	
	Pre- test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Shapiro-Wilk *p* value	.022	.011	.000	.000	.081	.051
Median	5.60	5.90	5.75	5.85	5.50	5.90
Range	3.60	3.60	3.30	3.50	2.00	2.50
Variance	0.52	0.48	0.54	0.70	0.79	0.91
Normal	No	No	No	No	Yes	Yes

Table 5.2 Comparison of pre- and post-test attainment scores (using the Wilcoxon signed ranks test) for the control and treatment group and the treatment group taught by the researcher and the control group

Ranks

Control_Treatment			N	Mean Rank	Sum of Ranks
Control	Post-test score – Pre_test score	Negative Ranks	13[a]	27.38	356.00
		Positive Ranks	57[b]	37.35	2129.00
		Ties	5[c]		
		Total	75		
Treatment	Post-test score – Pre_test score	Negative Ranks	22[a]	38.57	848.50
		Positive Ranks	50[b]	35.59	1779.50
		Ties	6[c]		
		Total	78		
Treatment (Res.)	Post-test score – Pre_test score	Negative Ranks	7[a]		
		Positive Ranks	19[b]	12.93 13.71	90.50 260.50
		Ties	1[c]		
		Total	27		

a.Post-test score < Pre_test score b. Post-test score > Pre_test score c. Post-test score = Pre_test score

Appendix 5B

Table 5.4 - Descriptive statistics comparing the change in attainment scores of the control and treatment groups, including a test to determine normal distribution

	Control Change *(post test score – pre test score)*	Treatment Change *(post test score – pre test score)*	Treatment Change(Res.) *(post test score – pre test score)*
Shapiro-Wilk *p* value	.220	.040	.241
Median	0.30	0.20	0.30
Range	2.00	2.30	2.1
Normal	Yes	No	Yes

Table 5.5 – Ranks produced by a Mann-Whitney U test comparing the control and treatment groups' mean change in attainment score

Ranks

	Control_ Treatment	N	Mean Rank	Sum of Ranks
Change_ Attainment	Control	75	86.0	6450.00
	Treatment	78	68.4	5331.00
	Total	153		

Table 5.7 – Ranks produced by a Mann-Whitney U test comparing the control and researcher-taught treatment groups' mean change in attainment score

Ranks

	Control_ Treatment(Res.)	N	Mean Rank	Sum of Ranks
Change_ Attainment	Control	75	52.48	3936.00
	Treatment(Res.)	27	48.78	1317.00
	Total	10 2		

Appendix 6A

Pupil Focus group transcript (Control)

Three pupils A (7b), K (7a), M (7b)

Question 1

R - First of all, have you noticed any changes in the way that you have been taught science during the school year so from the beginning of the school year to where you are now, have you noticed any changes? If you have what are they? And what did you think of them?

So do you want to start us off A

A- No

R – OK K?

K – I didn't notice any changes in the school year, I thought that it was taught quite consistently.

M – I thought that there were a lot of changes, we have been doing a lot more interactive, like more experiments and interacting instead of whatever is on the board we write it down and learn out of the textbook.

R – So at the beginning of the year were you just writing off the board?

M – A lot of the time and then it got more progressively more interactive

A – Well during our science lessons we have always kind of done practicals if we can but obviously if you learn something where you can't do one then we wouldn't, but I think ours has been the same, similar, because we've always done practicals. It has been different because we've had a different teacher and everything so it changed. One of them focussed on group work so we would get on working together, it was just about teamwork and making sure we knew what to do.

Question 2

R – Have a discussion with each other about how you have enjoyed or not enjoyed science whilst being taught specifically the micro-organisms, earth science, which is rocks and volcanoes, and magnets topics. So you will have done those quite

recently. So I want you to discuss how much you did or didn't enjoy it and give some reasons.

K – I really enjoyed the magnets one because we had a lot of practicals and tried to make electromagnets and tried to discover what magnets do and the field of magnetism or something like that. And we had quite a lot of practicals throughout in the earth sciences one we had rocks around the room that we had to try and figure out what type they were sedimentary, igneous or metamorphic and then write it down on a sheet, and we would discuss why we thought it was that.

R – So you enjoyed that?

K- Yep

M – I enjoyed them all, we did do a lot more practicals than we usually would have done in them three subjects. We did loads of practicals in the magnets one but I didn't enjoy it as much because it was more, I'm not sure, it didn't make me want to learn more, if you know what I mean. Although we did loads about it I think we did too much on the same thing, like week after week. But then we did like the volcanoes and the rocks it was always something different you we like learning something new every day and we did recap them but not as much as magnets which is what I thought made it a bit more boring.

R – What about microbes? Neither of you have spoken about microbes so far

K – I don't remember it

R – Don't worry it's not a test

A – I remember doing microbes, in all of them we did the same amount of practicals we normally would and for magnets I didn't really like it 'cause I found it boring. It wasn't my thing. But with microbes I enjoyed it I enjoyed looking at the history of it and I enjoyed the practical we did. We washed our hands with soap and washed them in just water and dried them and we all put a fingerprint on some slides and then we left them for a week or so and then we saw how much stuff, like bacteria, grew on the plate. It was really good and I enjoyed looking at smallpox and everything and like reading, I don't like reading from textbooks, but I enjoyed it because I enjoyed looking at the history behind it. I liked looking at the piece of paper as it had loads of cards about them and I had to try and figure out if it was a fungi, virus, or bacteria and it was really fun.

<u>Question 3</u>

R- Discuss whether, if you were asked that same question, about those three topics, would you have given the same answer before? In terms of enjoying science

before you did those three topics? Would you have given the same answer or a different answer and why?

A – I don't really know, it depends, I don't really enjoy science, I enjoy the practicals and the doing bit, like in chemistry, all the practicals, explosions and stuff and setting stuff on fire, I enjoy stuff like that. But I don't like remembering stuff and all the, to me, boring stuff, like memorising all these fancy words, and learning how to spell them and all that kafaffle. It's alright but its just not for me really.

R – So would you say that you have enjoyed those topics more or just the same as before?

A – Well, I kinda liked microbes more because it was more, I don't know, it wasn't necessarily the practicals, I find it more interesting, it was more down to Earth, it was a bit more casual, like you're not really going to find these fancy rocks are you unless you did it for a job but you know like stars. It was a bit more casual.

R – A bit more like everyday?

A – Yeh, its more everyday stuff and that you can actually and anyone could use it. Because if you had athletes foot, if I never learnt it then I more likely to use that information, just everyday knowledge really.

R – what about you guys, would you have given the same answer before these three units?

K – I would have probably given a similar answer 'cause I just enjoy science as a whole topic 'cause I find it interesting how things are related so like in chemistry they are all made out of like they look different but they are all made out of the same elements. And then with light, how everything reflects off everything even things that you wouldn't think are reflective, like your jumper. I just enjoy it as a whole because I find it really interesting.

M – I think that I would have given the same answer because I've studied a lot of the same stuff in primary school so I know what I find interesting and what I don't find interesting. But I also like learning, even if I don't think that I will use it when I'm older, I like to know in case I ever need the information.

<u>Question 4</u>

R – How effective, so how good, do you think that you have been taught the magnets, microbes and earth sciences topics has been to helping you learn? So just those three topics, how effective, how well do you think the way that you have been taught them has helped you learn?

A – I think that that it was alright, I think the way did help us to memorise stuff especially when it was leading up to the exams because it was good because they showed us different, normally teachers just show you in general one way of revising but in science my old science teacher showed us flashcards and everything or mind maps or you know just many, many ways of doing it and everything. I still don't remember it 'cause I'm not good at it. I feel that I've learnt and also our test on microbes we did, we had to write a leaflet or whatever to help someone remember if they were going on holiday, to help them. I liked that, I know it's a test, and I don't like tests but I found that it made it a bit more casual, and everything, and less panicky.

R – So they way that you were taught did you think that it was effective?

K – I think for earth sciences it definitely was because we had lots of examples of different types of rocks and what we were learning about, some videos, some slideshows there was some stuff on the board but there wasn't much. For the microbes it was alright because we had a few videos and some worksheets and then we had the agar plate experiment but it didn't really stick in my head as something that I enjoyed that much. I remember some of the information but I don't really remember how it links together. For the magnets I think that it was very well taught because I can remember how you can make an electromagnet better, what you can use to make an electromagnet, what the differences are between magnets and electromagnets and even what you use electromagnets for.

M – Obviously because we did so much recapping on magnets the information stays in your head more. But when we did the earth sciences one it was most of the time the way we had been taught so if was a slide show or we'd made a poster about it then you're gonna remember that because you have been thinking about it more. But for the microbes one we didn't do as many lessons on it so it wasn't as like drilled into your head as most other lessons, like the magnets. It wasn't really. It was like do the worksheet now watch the video, it wasn't really, OK the answers to the worksheet are this because so it's not like as imprinted as the other information.

<u>Question 5</u>

R – How focussed do you think you have been on learning when in science lessons since starting the microbes, magnets and earth science. How focussed have you been?

A – Not really; I have been more focussed on the microbes because I enjoyed it but not just 'cause I enjoyed it, I don't know, I found it interesting.

R – What if I said to you, what was you focus before you started those three topics?

A – Not really at all, I found it, I didn't find it fun, it was boring and I was like, sometimes I regret coming to science, I was like, urgh no its science. I know it helps if you enjoy it, it's alright.

R – So you wouldn't have said that your focus had changed apart from in the microbes topic?

A – It has a bit, but it's a bit like drowsy sometimes, it's really quiet and boring.

R – K, Focus during those three topics?

K – I think I was quite focussed in those three topics as I didn't really know anything about the earth science or magnets, I knew a bit about the microbes. But I learnt quite a few new things in earth science and magnets.

R – And what about your focus before?

K – Before I think that I was still quite focussed because I think I had a good balance of enjoying myself doing the topics I liked and focussing on the information and remembering everything.

M- I have always focussed in science, because I really enjoy science so I don't think my focus has changed. But even if I didn't enjoy it as much as a topic I wouldn't not focus because I still need to know this information as its still being taught to me for a reason, so I'm always going to listen and get all the information I need.

<u>Question 6</u>

R- How good have you been at completing work set. So in class how good have you been at completing it since you started the microbes, magnets and earth science topics?

A – I don't know, if I've learnt it or know it then I find it easier but then if I don't understand something then I find it hard to complete it cause with some of the stuff that I don't understand it I get confused and I sometimes don't complete stuff, because I don't understand it.

R – And has that changed just in those three topics to previous or would you say that you've been the same all the way through

A – It's been the same as I don't really understand science really

K – I think I've been very good at completing work because I have to get lots of extensions and I'm usually finishing my work and then like helping out people with their work.

R- And has this changed in these topics compared to other topics, or are you always the same?

K – I think I'm always the same.

M – I always try to finish my work, I always have done, as if you're going to go into exams you need to revise. Oh no I haven't finished this piece of work I don't know the information it could be one mark difference in getting that B or A. So I always complete my work or at least try to, and even if I don't I'd try to do it at home or ask sir.

R – Do you think that that has changed throughout the year or are you always like that?

M – No, I've always done that.

Appendix 6B

Pupil Focus group transcript (Treatment)

Six pupils (KS3 target level) C (L7b), Ch (L7a), L (7a), F (7b), J (7b), D (7b)

Question 1

R – Have you noticed any changes in the way that you have been taught science during this school year? If you have what are they? And what did you think of them?

C – We started working a lot more with groups with other people and sort of helping them learn instead of just independently learning.

R – What did you think of that?

C – I like it, I like working in groups, I like sharing my ideas and being a bit conversational.

Ch – Like C said, we started working in groups but I think that sometimes then one person in the group that always did the work instead of everyone else chipping in they just let one person do it and then they just like chilled and didn't do very much.

F – Well we are doing lots of big projects that we are working together on which I think is good because then we can all help each other and people who are at different levels that might not understand things like however other people might not be so dedicated to work so maybe they could be encouraged to help out in the team.

J – We have been doing more group work for projects of about two or three weeks, I hardly seem to enjoy it as much as I did when we had our other teacher, we did a lot more practical and single work which I found more easier then working in a group, where you have two people doing work and everyone else just laying back.

L – We have been working in groups and I haven't enjoyed because it's been a lot of practical work which has taken a very very long time to plan because people haven't been focussing so I feel it's been very one sided from what we could have learned individually, a lot of people made no progress, other people almost went backwards because they were trying to do too much and trying to compensate for things that they shouldn't of had to compensate for.

R – Dominic what do you think?

D – I don’t know the question

R – The question was, have you noticed any changes in the way that you have been taught science this year, if you have what are they? And what do you think of them?

D – I haven’t noticed anything

Question 2

R – So have a discussion with each other, remember to say your name, about how much you have enjoyed or not enjoyed science whilst being taught the microbes, Earth science, which is rocks, and magnets topics and give some reasons why you have or haven’t enjoyed those topics.

C – I liked Earth science because it was quite interesting, I learned things that I hadn’t learnt before like the acids and alkalis attacking rocks and we got to do a lot of practicals and do a carousel. The microbes was good because the yeast experiment was fun and it showed you more about what microbes did and how much they are actually in. Magnets was a bit predictable and we put too many practicals into one lesson, it was rushed and you didn’t have enough time to write down the answers.

R – What do other people think, want to add or take away from that?

L – I didn’t enjoy it at all. We did a lot of practicals and they were rushed and what we could have learnt without the practicals would actually have been more productive because we would have had time to go into detail and explain it in a different way rather than spend quite a while trying to prove a fact we know to be true already.

R – Dominic? (silence) no thoughts? (silence) So you don’t think that things have changed? (silence) No? OK.

J – I’ve not really enjoyed it. I’ve enjoyed the fact that it has been a lot of practicals but we’ve been put into groups that we’re not really happy with, it’s been people who have been randomly selected to go into groups, and we have people who are actually working and trying to do something and get good marks while others are laying back at letting everyone else do the experiments.

Ch – I think that I didn’t really enjoy the magnets one because like everyone said, it was quite rushed and we didn’t really take out time to get the best outcome of the experiment. And I enjoyed the one, the earth science, because, well, when we were doing it in my group there was only two of us and we were the people who actually did the work together. So when the boys came back they didn’t really do much anyway so they kind of just copied us. And then for the one where we had the yeast, microbes, I think that, that practical was good ‘cause we got to choose what practical we wanted to do out of a booklet of them, then see if we could get the best outcome.

F – I enjoyed all the topics and although we did work in teams that we weren’t happy with I think we got a lot out of it by using stations and going round and

finding out information. At the end of each topic we did a write up and also found out some new information to come to a conclusion of what we had learned.

<u>Question 3</u>

R – Going to come to you first Dominic. Right I asked people questions about whether they enjoyed the last three topics, if I asked you the same question about the topics previous to those three topics what would you have said about whether you’d enjoyed them or not?

D – Can’t remember the topics

Ch – I enjoyed doing the respiration one because we were doing about the lungs and stuff and then like how smoking affected it, and like asthma and stuff, and because I have asthma it kind of like opened my eyes a bit to what was actually happening inside my lungs when I start having an asthma attack ‘cause I didn’t really understand it. That was good, and we got to split open a pig’s lung.

C – Like Charlotte said, I did like the bit where we dissected a lung, not because it was gory and you know, it was actually informative. We had seen diagrams of the tubes that the air went down but when we dissected the lung we could actually see which were behind which, where it was, how long it was, it was really accurate, I actually learnt a lot even if I did feel a bit queasy.

J – I enjoyed pretty much all the topics previous because Mrs X made all the lessons really fun. Sometimes it was written, then there was a lot of practicals in it as well and that made me feel really comfortable because I enjoy practicals more than theory.

L – I liked it because we had quite a bit of theory so what we learnt, we learnt in detail in a lesson where we could link up all the information while it was still fresh in our minds. And because it was written we had concise revision notes rather than diagrams or tables where we had to still try and remember what we were doing exactly.

J – I find that when you do activities, like how you did the heart and everything, you got involved in it, you remembered it more rather than just copying from a textbook and it gets you thinking about what it actually is and why it would happen.

R – So do you think that the way that you have enjoyed science has changed through the year; have you enjoyed it more recently or less recently?

C – I've enjoyed it less recently because, I like working in teams but sometimes we were put in quite bad teams and we did spend an extraordinary amount of time planning and sort of going over the same answer twice so I think I liked earlier on when we just did the practicals and then evaluated it after we had actually got the knowledge into our heads.

Ch – I enjoyed it like, not too much less, just a little bit less when we were in the groups I was one of the people who did most of the work and I tried to get everyone else to help but they were just like, not doing anything. I didn't feel like it was fair to me to be doing all the work and then just lounging about.

J – I've enjoyed it a lot less because I feel really pressured by the rest of the group. 'Cause we get put into groups where you have very low level science and very high level science, and the lower level expect all the higher level to do all the work, they don't want to do anything.

C – I think next year when we get put into sets it will be better, people will be more dedicated to work, and others, that might not be a good thing actually as people who aren't that keen on work, they probably wouldn't be influenced by other people to show them what the right thing to do and help them learn but I think that it would be better for me personally and others who want to learn, to actually get on with it and make good progress.

<u>Question 4</u>

R – How effective do you think that the way that you have been taught, the micro-organisms, magnets and earth science topic has been in helping you learn? D?

D – I think it's been OK in helping me learn

F – I think that by doing these recent topics it's more fresh in my memory and we've covered it quite a bit more so, and its helped me learn because we get involved more with those topics.

C – The earth science has particularly stuck in my mind especially when we learnt about the mantle and the core of the Earth because I'm quite a, I learn from pictures. When we did the big posters and diagrams of the Earth I actually remembered that because I've got it sort of labelled in my mind so I think that having to make a lot of posters from your memory and from scratch was quite helpful because you could picture what you had written, not what somebody else has written in the text book.

Ch – I agree with C, 'cause I think that I learn more from pictures as well, rather than writing stuff down, and that it would be easier for making posters to make sure that it stays in your mind then you can go into more detail to find out more.

J – I'm quite different from you, I learn more through practical, getting my hands in stuff and there were a few practicals in the topic and I was happy with that but I feel that I only remember as much as the other lessons.

L – I didn't, I feel that it wasn't as good for my learning because I've found that I been struggling to get things to work efficiently so there hasn't been time to do what I like to do to memorise it and to make my own links and knowledge. I haven't had time to do that as I have been trying to make sure that the basic facts are down for everybody.

<u>Question 5</u>

R – How focussed do you think that you have been learning the earth science and magnets and microbes topics? So how focussed do you think that you have been? The follow on question for that is; do you think that it has changed throughout the year? Your personal focus.

F – I think that I have quite good focus because I am quite keen to learn but, and I think that these topics have been quite interesting, so I think that topics that are quite interesting you are more focussed on so maybe if all topics are maybe made more interesting by maybe more practicals then or maybe things that people like it would help people keep more focussed.

R – Do you think that your focus has changed this year or that you've just been the same?

F – Well maybe maturity, like everyone, focus is increased.

Ch – I think that this year I, at the beginning of the year I came in really focussed and then towards the end of the year when we've been put into groups I may have lost my focus a bit because other people around me who are messing about a bit are making me laugh, puts my focus a bit wobbly. But then I think that sometimes I think that I prefer doing it just in partners like where we're sat so that we're away, well not exactly away from everyone else but so that we're more focussed together rather than other people coming in having a look at what we're doing and going out and going back to their groups.

C – I agree with Ch, for sitting in pairs is better than sitting in groups. When you're in a pair it's a sort of informed discussion when you're in a group it's more of a debate. It's like there's some people who actually know the facts and are trying to get it across to people with the less science knowledge and people who are convinced that they're, that it's like the opposite opinion so it's more of a who can shout the loudest and who knows the most that they can yell at the other person.

R – So what affect do you think that that has had on your focus?

C – I think that in a pair you've got better focus, so in the beginning of the year I was probably more focussed mainly because my memory was fresh from the summer holidays, it was a new year an' all but as we got into groups, like Ch said, other people who don't really know what they are doing are distract you if you try to help them instead of helping your memory and mind.

J – I agree with C and Ch, I think my concentration rate dropped quite a bit because the people in my group there are some of us, like I said earlier, that have been trying to do something then the others just being lazy and messing around so the people who are actually wanting to do something are having to wait until the others are stopped or trying to say you need to stop now which means we all get bored of waiting and our concentration just drops.

R – Your focus D through this year?

D – My focus has been OK on the subjects that are interesting but the ones that aren't so it's been not very focussed on some of them.

L – My focus this year, the levels and concentration span have been just as high but what I'm focussing on has been different because we can be two or three tasks behind in our group and I may have finished but we've got to get the rest of the group to the same place so I could be told that we need to move on to something else and while trying to remember that and the information I need I also need to get the rest of the group to focus on what they needed to do so that they can get there.

<u>Question 6</u>

R – How good have you been at completing the work that you have been set in those topics; microbes, magnets and earth science and do you think your ability to complete work set has changed during the year?

C – I don't think it's changed at all I'm still motivated, I still want to learn, you never really stop learning and even though we were in groups and it sort of distracts you from the actual classwork, when it's homework and it's the sort of it's the conclusion of the subject my focus hasn't dropped, it's still the same.

R – So you're still getting your work completed?

C – Yes it hasn't changed that

F – I'm quite an independent person so I always work by myself and get work completed so if there's everyone else around me, sometimes they're not, they're quite slow and it sort of makes you, I don't know, if everyone else is chatting you try to just get on with it, I just do.

Ch – I find that because we've been in groups for the last few topics we've been trying to finish, complete the work together but then some people don't chip in their ideas and like I said before some people leave like one person to do all the work but then sometimes they don't really know the answer so you try and get the others to help you, but then they don't know the answer as well so they don't really help so then it's quite hard to finish it. I have found that , that has been quite the case with my group

J – I think my ability to answer questions has gotten better but my ability to explain isn't as good because we've just been left in our groups the teacher says OK gonna give you a subject then go round in the practicals, you haven't been given much explanation about it.

R – So do you think that your ability to complete work has changed during the year? Or for you have you stayed the same?

J – I think about the same

R – What do you think D? Do you think your ability to complete work has changed at all this year?

D – Not really I have completed what I need to complete and if I haven't then I have done it out of lesson.

L – Completing work in lessons has gone downhill at one point I had to change experiment and I actually got half way through that experiment before anyone in my group actually realised that I'd changed experiment despite the fact that I had told them three times the experiment has changed. But I'm trying, the level of motivation for homework hasn't changed so that is where I am trying to make up for it, I'm trying to catch up with the work that I am doing.

Appendix 6C
Pupil Focus Group - Control and Treatment thematic comparison

Colour	Theme (mentions) - Control	Quotes - Control	Theme (mentions) - Treatment	Quotes – Treatment
Green	**Positive Factors** • Practical work (5) • Topic subject (6) • Discovery/ analysis (3) • Teaching strategies (12)	"I really enjoyed the magnets one because we had a lot of practicals" "But with microbes I enjoyed it… It was really good and I enjoyed looking at smallpox" "…we had to try and figure out what type they were sedimentary, igneous or metamorphic and then write it down on a sheet…" "...examples of different types of rocks… some videos, some slideshows there was some stuff on the board but there wasn't much."	**Positive factors** • Practical work (7) • Topic subject (5) • Discovery/ analysis (5) • Teaching strategies (5) • Group work (3)	"I learned things that I hadn't learnt before like the acids and alkalis attacking rocks and we got to do a lot of practicals and do a carousel." "The microbes was good because the yeast experiment was fun and it showed you more about what microbes did…" "… I think we got a lot out of it by using stations and going round and finding out information. At the end of each topic we did a write up…" "When we did the big posters and diagrams of the Earth I actually remembered that because I've got it sort of labelled in my mind…" "…, I like working in groups, I like sharing my ideas and being a bit conversational."
Red	**Negative factors** • Practical work (2) • Topic subject (3) • Teaching strategies (1)	"… practicals in the magnets one but I didn't enjoy it as much because it was more, I'm not sure, it didn't make me want to learn more…" "… I didn't find it fun, it was boring and I was like, sometimes I regret coming to science…" "It was like do the	**Negative factors** • Practical work (4) • Topic subject (2) • Teaching strategies (2)	". Magnets was a bit predictable and we put too many practicals into one lesson, it was rushed…" "… that I didn't really enjoy the magnets one because like everyone said, it was quite rushed." "… so there hasn't been time to do what I like to do to memorise it and to make my own

		worksheet now watch the video, it wasn't really, OK the answers to the worksheet are this because so it's not like as imprinted as the other information."	• Group work (11)	links and knowledge." "'Cause we get put into groups where you have very low level science and very high level science, and the lower level expect all the higher level to do all the work…"
Yellow	**Recognition of teaching strategies** • Practicals (10) Positive (5) Negative (2) Neutral (3) • Group work (1) Neutral (1) • Discussion (1) Positive (1) • Textbooks (3) Positive (2) Neutral (1) • Slideshow (2) Positive (2) • Investigating (3) Positive (3) • Flashcards (1) Positive(1) • Mindmaps (1) Positive (1) • Videos (3) Positive (2) Neutral (1) • Posters (1) Positive (1)	"I don't like reading from textbooks, but I enjoyed it because I enjoyed looking at the history behind it. I liked looking at the piece of paper as it had loads of cards about them and I had to try and figure out if it was a fungi, virus, or bacteria…" "…we had quite a lot of practicals throughout in the earth sciences one we had rocks around the room that we had to try and figure out what type they were… we would discuss why we thought it was that." "…I think the way did help us to memorise stuff especially when it was leading up to the exams…[my]teacher showed us flashcards and everything or mind maps or you know just many, many ways of doing it…"	**Recognition of teaching strategies** • Practicals (11) Positive (7) Negative (4) • Group work (15) Positive (3) Negative (11) Neutral (1) • Textbooks (2) Negative (2) • Investigating/planning(5) Positive (3) Negative(2) • Posters (2) Positive (2)	" …working in groups but I think that sometimes then one person in the group that always did the work instead of everyone else chipping in." "…when you do activities, like how you did the heart and everything, you got involved in it, you remembered it more rather than just copying from a textbook…" "…we did spend an extraordinary amount of time planning and sort of going over the same answer twice so I think I liked earlier on when we just did the practicals and then evaluated…" "…When we did the big posters and diagrams of the Earth… so I think that having to make a lot of posters from your memory and from scratch was quite helpful because you could picture what you had written, not what somebody else has written in the text book."

	• Worksheets (2) Positive(1) Neutral (1) • Board work (2) Neutral (2) • Leaflet (1) Positive (1) • Tests (1) Negative (1)			
Pink	**Work ethic** • Focus positive (5) negative(2) • Completion of work positive (2) Negative (1)	"… still quite focussed because I think I had a good balance of enjoying myself doing the topics." " Not really at all, I found it, I didn't find it fun, it was boring…" "…the stuff that I don't understand it I get confused and I sometimes don't complete stuff, because I don't understand it." "I've been very good at completing work because I have to get lots of extensions and I'm usually finishing my work and then like helping out people with their work."	**Work ethic** • Focus (17) Positive (8) Negative (6) Neutral (3) • Completion of work (5) Positive (4) Negative (1)	"…I'm still motivated, I still want to learn … subject my focus hasn't dropped, it's still the same." "… the levels and concentration span have been just as high but what I'm focussing on has been different because we can be two or three tasks behind in our group…" "Completing work in lessons has gone downhill…, the level of motivation for homework hasn't changed so that is where I am trying to make up for it…" "…trying to finish, complete the work together but then some people… don't really help so then it's quite hard to finish it"

Appendix 6D

Teacher Focus group transcript

Four teachers P, C, L, A

Question 1

R – P do you want to start us with how have you found the teaching and delivery of the inquiry-based schemes of work provided?

P - Well Emma, the schemes were fine to follow, very straight forward, the one, the two little bits that threw me at first because it was all set out very clearly and logically I didn't do as much preparation as I should have done to think about where the learning was going in the lesson to do the intervention with the small groups. Once or twice I missed doing all the copying until last minute so it was a bit of a panic to get all the copying ready for the next lesson but that was the only minor quibble.

L – Well, I enjoyed teaching it but there was a lot to fit in each lesson though and that's where I struggled, fitting everything in. Some of the materials probably could have been simplified some of it was probably a bit too complicated for year 8, but other than that it was good to teach and very straight forward to teach.

C – Yep, planning wise it was all there, all the resources are there. I only did the magnetism one and I think the first two lessons, it was all those practicals, maybe that could have been split into, it could have been into two separate lessons, I only saw them once a week and it ran on too long and they had kind of forgotten what they had done when we came to talk about it but yes it was fairly easy to teach as all the stuff was there.

A – Yes, I enjoyed doing the bit, it's nice getting the kids to look stuff up by themselves. It needs a lot of thought when you are preparing stuff about what sort of questions you are going to ask the kids to lead them on without giving them the answer, because this is meant to be about them finding stuff out for themselves, and I found personally that that was the most difficult thing, is getting that level right.

P – With my group, I had a really good TA who was very good at getting the kids to talk to each other and come up with a set of questions for me to then go through with them which were science related on the topics that they were doing, she was really good at focussing them in on it. Where it fell down a little bit was in the groups where they didn't all engage and they allowed the others not to partake.

A – Can I speak up on that, with the groups where you haven't got everyone partaking I think that the kids need a lot more training in terms of how to interact with the information, how the people in the group must deal with someone who isn't helping and isn't on board and that needs a bit more training up, it's something that they have got to learn to do.

Question 2

R – The next question is, this a bit more of a general sort of question about the style the strategy of teaching in this way. So how valuable do you feel that this style, this inquiry-based learning style is to pupils?

P – I think the understanding is as good as it is with any other style, some of the students that got involved were much more engaging with the questions and coming up with more connections and links between things. It wasn't right across the board but some of the class were doing that. What I found was, I don't know if it was the style, or the way I taught it but having the summary facts, they didn't feel that they had that scaffold to put everything on at the end when typically the brighter ones come back and review it and they missed that a bit I think.

R – L what do you think about the value of this style of teaching to the learning of pupils?

L- I think that always in any style of teaching there are those that really enjoy and thrive on it, I don't know if this style of teaching it makes it easier for some people to coast or just to be, yeah, not participate because others will do the work for them. But I think that as they had to think for themselves a lot, those who were engaging really knew what they were doing rather than just following a teacher telling them what to do.

C – Yes definitely a place or it, I wasn't convinced about doing it as a whole big series of lessons. I felt that by the end of 6, 7, 8 lessons that some of the weaker pupils had kinda got a bit lost and maybe that was my fault for not picking up on it but it was easier for some of them to coast, easier for ones in a group to not really do very much. So maybe introduce a topic more traditionally, teacher-led perhaps, and then have a few lessons like this and then sort of tie it up at the end. I felt by the end of the topic some of them I didn't feel had learnt very much, some of them had but some of them had sort of coasted along and not…..

P – I wonder how much they would normally do that in a lesson anyway, it's hard to judge. That's the bit that I find hard to judge, it's the comparison.

C- Yeah, yeah

R – What do you think A about this style of teaching and its value?

A – Different kids enjoyed it more than they would have done in the standard style and I've found a number of kids do get a lot out of this style. When we are focussed on exams we do have to think about what sort of information is the kid actually taking away and I think that you probably do need to have some sort of scaffolding for them to bring the ideas together at the end, and you should introduce the scaffold beforehand so that when they are doing all of the different activities, which I think gets them thinking, then they can bring that together and think, what have we learned? How would I arrange this? You know, How do my ideas pan out? Oh I've found this, that links with this; building it all up. Now I like this style but it does need a lot of, initially, a lot of investment of time to get that, what do we actually want at the end of this? So they've got a picture. Does that make sense?

<u>Question 3</u>

R – OK so the last one, I'm going to give you a few seconds to read what my research questions are. So obviously with your experience of teaching in general and your experience of teaching this style, what do you think your thoughts are on the answers to these? If you were to give any sort of ideas for what you think?

P – I think I'm sort of going with the flow of teachers here that think it's got its place but not as the only method of teaching. I think if we can get the kids engaged in group work, it's going to increase independence, I saw a lot of signs of that going through the topics and it's going to get students more involved in delving into ideas that interest them, exploring them and how you explore topics scientifically, I think enthusiasm.

R – So that's kind of like perception

P – Perception, yes. In terms of the attainment I couldn't see any difference this way or the other way. It appealed to different groups of students, I had a few bright students who rely on the structure and their books to revise at the end who didn't like it, and others who don't normally do particularly well engaged with it and did better. So I think we had a mixture there. Critical thinking skills, worked well and I think repeated over a longer period of time then they can only get better as they are practising using those ideas more usefully rather than a one off practical which is what we normally do, so that critical thinking skills and developing independent learning, I think is where it is going to be most useful.

L – Just what P said really, yeah attainment, I couldn't really tell whether there would be any difference there. The critical thinking skills, that is probably where it's most useful, really making them think about what they are doing. I think the inquiry-based learning will appeal to different pupils and for some of them, it will really help them enjoy science lessons more than traditional teaching. At the same time there is those students who, I would be one of them, who like the traditional lesson.

C – Attainment, kind of hard to measure, magnetism questions on the test. I taught one group that was doing this scheme and one that wasn't, there wasn't a huge amount of difference. Critical thinking, they were starting to develop that, I think that is something we need to work on. I felt in some groups there would be one or two pupils who would be kind of doing all the thinking for the rest of the group so it's how to train them all to think and not just to sit back. Then perception, some of them enjoyed it, some of them didn't enjoy it so much and I think maybe mixing and matching a bit so that it wasn't so 'oh we're doing this again' for the ones that didn't enjoy it. Yeah if you did a couple of lessons and then did something different but I think its good, its good to have a different style.

A – I agree with what has been said ahead but a couple of things. I think some of the times the kids got information overload, there was so much information they couldn't decide just which were the important points to pull things together. That, I think, just comes back to training though and the more they do that the better they get at it. On the back of that, what is the best question for them to be asking? They could be asking lots of different questions but it wasn't necessarily taking them forwards in terms of moving their learning on, getting to grips with something that was useful , shall we say. It's that thing again, because of lack of training, it's a good thing to do but it does take a lot of time for the kids to get this idea of; they're responsible for the learning, they're responsible for taking out the information and assessing it, deciding what's important and what's not. Then using the important stuff to answer the original question, and that takes a lot of practice. Pupil attainment seems similar to what it was before, some were better some found it not as easy to learn from, possibly because we didn't get to Here are your key points and bullet points. Critical thinking, I think that's what I was answering in the first bit. Perception of science, some loved it some didn't same as anything, it doesn't matter what you do there will always be some that like it and some that don't, so mix and match.

Colour	Theme (mentions) - Teacher	Quotes - Teacher	Similar theme (mentions) - Pupil	Quotes – Pupil
Green	**Positive Factors** • Prepared resources (3) • Enjoyment of teaching(3) • Pupil enjoyment (12) • Valuable teaching style (13)	"…the schemes were fine to follow, very straight forward…" "Now I like this style…" "Well, I enjoyed teaching it…" "Different kids enjoyed it more than they would have done in the standard style and I've found a number of kids do get a lot out of this style." " it's a good thing to do but it does take a lot of time for the kids to get this idea of; they're responsible for the learning" , so that critical thinking skills and developing independent learning, I think is where it is going to be most useful. ".. and for some of them, it will really help them enjoy science lessons more than traditional teaching."	**Positive factors** • Practical work (7) • Topic subject (5) Valuable teaching style • Discovery/analysis (5) • Teaching strategies (5) • Group work (3)	Not mentioned by teachers Not mentioned by teachers "… I think we got a lot out of it by using stations and going round and finding out information. At the end of each topic we did a write up…" "When we did the big posters and diagrams of the Earth I actually remembered that because I've got it sort of labelled in my mind…" "…, I like working in groups, I like sharing my ideas and being a bit conversational."
Red	**Negative factors** • Teacher	" I didn't do as much preparation as I	**Negative factors** • Topic subject (2)	Not mentioned by teachers

	preparation (4) • Group work (5) • Key points written down (4) • Volume of information (5)	should have done to think about where the learning was going in the lesson to do the intervention with the small groups" "…but it was easier for some of them to coast, easier for ones in a group to not really do very much…" "…some found it not as easy to learn from, possibly because we didn't get to Here are your key points and bullet points…" " I think some of the times the kids got information overload, there was so much information they couldn't decide just which were the important points to pull things together." "… the magnetism one and I think the first two lessons, it was all those practicals, maybe that could have been split into, it could have been into two separate lessons…"	• Group work (11) • Teaching strategies (2) • Practical work (4)	"'Cause we get put into groups where you have very low level science and very high level science, and the lower level expect all the higher level to do all the work…" "… so there hasn't been time to do what I like to do to memorise it and to make my own links and knowledge." ". Magnets was a bit predictable and we put too many practicals into one lesson, it was rushed…"
Yellow	**Improvement suggestio**	"… kids need a lot more training	**Recognition of teaching strategies** • Practicals (11)	" …working in groups but I think that

	ns • Training in group work (3) • Training on analysing information (2) • Teachers preparing for questions to scaffold (3) • Summary of facts for test at the end (5) • Mix and match strategies (6) • Use of	in terms of how to interact with the information, how the people in the group must deal with someone who isn't helping and isn't on board…" ". It needs a lot of thought when you are preparing stuff about what sort of questions you are going to ask the kids to lead them on without giving them the answer…" "When we are focussed on exams we do have to think about what sort of information is the kid actually taking away and I think that you probably do need to have some sort of scaffolding for them to bring the ideas together at the end." "it doesn't matter what you do there will always be some that like it and some that don't, so mix and match." "…I had a really good TA who was very good at getting the kids to talk to each other and come up	7 positive 4 negative • Group work (15) 3 positive 11 negative 1 neutral • Textbooks (2) 2 negative • Investigating/planning(5 3 positive 2 negative • Posters (2) 2 positive	sometimes then one person in the group that always did the work instead of everyone else chipping in." "…when you do activities, like how you did the heart and everything, you got involved in it, you remembered it more rather than just copying from a textbook…" "…we did spend an extraordinary amount of time planning and sort of going over the same answer twice so I think I liked earlier on when we just did the practicals and then evaluated…" "…When we did the big posters and diagrams of the Earth… so I think that having to make a lot of posters from your memory and from scratch was quite helpful because you could picture what you had written, not what somebody else has written in the text book."

	TA (1)	with a set of questions for me to then go through with them…”		
Pink	**Attainment** • Similar (4) • Different groups (5)	“ I think the understanding is as good as it is with any other style…” “…attainment I couldn't see any difference this way or the other way. It appealed to different groups of students, I had a few bright students who rely on the structure and their books to revise at the end who didn't like it, and others who don't normally do particularly well engaged with it and did better.”	**Completing work** • Completion of work (5) Positive (4) Negative (1)	“Completing work in lessons has gone downhill…, the level of motivation for homework hasn't changed so that is where I am trying to make up for it…” “…trying to finish, complete the work together but then some people… don't really help so then it's quite hard to finish it” “… the levels and concentration span have been just as high but what I'm focussing on has been different because we can be two or three tasks behind in our group…”
Blue	**Critical thinking skills** • Improvement (4) • Types (4)	“.. Critical thinking skills, worked well and I think repeated over a longer period of time then they can only get better as they are practising using those ideas more usefully.” “…they're responsible for the learning, they're responsible for taking out the information and assessing it, deciding what's important and	**Teaching strategy comments - no mention of skills from pupils**	“… I think we got a lot out of it by using stations and going round and finding out information. At the end of each topic we did a write up…” “When we did the big posters and diagrams of the Earth I actually remembered that because I've got it sort of labelled in my mind…” “…, I like working in groups, I like sharing my ideas and being a bit conversational.” “…we did spend an extraordinary amount of time planning and

	• Training (5)	what's not. Then using the important stuff to answer the original question, and that takes a lot of practice…" "I think if we can get the kids engaged in group work, it's going to increase independence, I saw a lot of signs of that going through the topics." "… kids need a lot more training in terms of how to interact with the information… They could be asking lots of different questions but it wasn't necessarily taking them forwards in terms of moving their learning on."		sort of going over the same answer twice so I think I liked earlier on when we just did the practicals and then evaluated…" "…trying to finish, complete the work together but then some people… don't really help so then it's quite hard to finish it" "'Cause we get put into groups where you have very low level science and very high level science, and the lower level expect all the higher level to do all the work…"
Grey	**Pupil perception (22)** Positive (12) Negative (10)	"… some loved it some didn't same as anything.." "… will appeal to different pupils and for some of them, it will really help them enjoy science lessons more than traditional teaching. At the same time there is those students who, I would be one of them, who like the traditional lesson."	**Pupil perception** • Positive (25) 7 Practical and 5 topic which weren't mentioned by teachers and are not a unique feature of IBL • Negative (19) 11 of these were about group work	"… I think we got a lot out of it by using stations and going round and finding out information. At the end of each topic we did a write up…" "When we did the big posters and diagrams of the Earth I actually remembered that because I've got it sort of labelled in my mind…" "…, I like working in groups, I like sharing my ideas and being a bit conversational." "…we did spend an extraordinary amount

				of time planning and sort of going over the same answer twice so I think I liked earlier on when we just did the practicals and then evaluated…" "'Cause we get put into groups where you have very low level science and very high level science, and the lower level expect all the higher level to do all the work…"

Appendix 6F

Table 6.1 Qualitative Summary of Treatment and Control Groups' Answers to Questions

Control summary	Treatment summary	Comparison
1. Have you noticed any changes in the way that you have been taught science during this school year, if you have what are they and what did you think of them?		
Two (67%) noticed no changes except for differences between teachers; one (33%) stated that there were differences. They recognised differences in teaching styles, mentioning interactive, practical, group work and copying off the board and/or working from textbooks.	Five (83%) noticed that they had been working in groups more. Four (67%) stated that they were helping others to learn. 50% of the pupils were positive and 50% negative about the change.	Pupils in the treatment group noticed a change in teaching strategy compared to the control group. Control group mentioned a variety of strategies that were used whereas the treatment group focussed on the group work element.
2. Have a discussion with each other about how much you have enjoyed or not enjoyed science whilst being taught the microbes, earth science and magnets topics and give some reasons		
All pupils (100%) said that they enjoyed practical work and secondly that they enjoyed working things out, such as identifying rocks or diseases from information. Although they all enjoyed the practical they each favoured one of the three topics and negative points focussed on the subject enjoyment rather than teaching style. One pupil (33%) thought that there was the same amount of practical work as usual and the other two (67%)	Five pupils (83%) mentioned two factors: practical work and group work. 50% enjoyed the practical work and 50% didn't. The ones that did not stated it was because they felt they were rushed and were spending too long helping others or doing the work for others who were not engaged. Positive comments included enjoyment of moving between stations and having choice in	Both groups enjoyed practical work and/or finding information and working something out, although in the treatment group positive feelings towards all teaching strategies were affected by working in groups. The treatment group focussed on the teaching style whilst the control group it was on topic content.

implied that there was more in at least two of the topics.	investigations.	
3. Discuss whether if you were asked this same question about the previous topics whether you would have given the same answer? Why?		
All pupils (100%) would have given the same answers. One (33%) found microbes subject more interesting but in general was not keen on science, the other two (66%) enjoyed science and their enjoyment had not changed.	All pupils (100%) enjoyed the previous topics more, although they focused on respiration. They enjoyed the dissection, practical work and having concise notes. They did not like the treatment lessons as they felt they had to do the work for the less able pupils in their groups.	In the control group enjoyment of science remained steady. In the treatment group they had preferred topics previously due, in part, to there being less group work, although only mentioning one previous topic.
4. How effective do you think the way that you have been taught the microbes, magnets and earth science topics has been to helping you learn?		
One pupil (33%) focussed on revision techniques and assessments that helped her. Two (67%) discussed teaching strategies video i.e. clips, worksheets and reduced writing. One (33%) stated that the repetition of similar points in magnets was good; the other thought that it was boring. This linked to their preference for the topic.	Five pupils (83%) felt that they had learned effectively and found practicals, posters and investigations a good way to learn. One of the pupils (17%) felt that she did not have enough time to make her own notes and consolidate understanding.	The control group pupils talked about a variety of teaching strategies largely in the positive. The treatment group were largely positive about how much they had learned although one pupil felt that the teaching style had a negative effect on her learning.

5. **How focussed do you think you have been on learning when in science lessons since starting the microbes, magnets and earth science topics?** ▫ **Do you think your level of focus has changed through the year? Why?**		
All pupils (100%) said that their focus hadn't changed. Two (66%) stated that they are generally always focussed. The other pupil said that she doesn't like science so doesn't focus except in the microbes topic which she found interesting.	Four pupils (67%) felt that their focus had dropped due to group work, as they had to debate more and were more distracted. Two pupils (33%) stated that their focus had increased on different things such as the group work dynamic. One pupil (17%) said that it was topic-dependent. All pupils (100%) said that pairs are the optimum group size for focus.	Pupils in both groups stated that their focus was fairly consistent. In each group one pupil said that their focus was topic-dependant. In the treatment group some thought that their focus was negatively affected by group work.
5. **How good have you been at completing work set since starting the microbes, magnets and earth science topics?** ▫ **Do you think that this has changed throughout the year? Why?**		
All pupils (100%) agreed that their work completion had not changed throughout the year.	Five pupils (83%) felt that they had continued to complete all work but two pupils (33%) thought that they were completing less in class due to having to group work, but that they would complete it at home.	In both groups pupils stated that work compleetion had not altered but in the treatment group two pupils felt that they were getting less completed in class time.

Appendix 6G

Table 6.3 Qualitative Summary of Teachers' Answers to Questions Compared

Positive	Negative
1. How have you found the teaching and delivery of the inquiry-based schemes of work provided?	
	50% thought there was too much in the lessons
100% said that resources and planning were easy to follow.	50% felt that they should have better planned questioning and scaffolding in advance
75% enjoyed teaching it.	50% mentioned that group work allowed some pupils not to engage but it was commented that this could be improved with pupil training.
2. How valuable do you feel this inquiry-based learning style is to pupils?	
100% felt that there was no difference overall in pupils learning of subject knowledge, between inquiry-based and their usual teaching style.	
100% stated their belief that critical thinking skills were improved and 50% gave examples of the types of skills.	50% made comments on pupils requiring training to make these strategies more effective.
100% believed that this style of teaching should be incorporated into teaching alongside other types to provide variety.	75% of teachers felt that they had not been enabled to tie up key points for exam preparation at the end of the topic.

3. Read my research areas. What are your thoughts on these? • *The impact of inquiry-based teaching strategies on pupils' critical thinking skills* • *The impact of inquiry-based teaching strategies on pupils' attainment* • *Pupils and teachers perception of inquiry-based teaching strategies*	
100% felt that it was positive for critical thinking skills and engagement for some pupils. The skills mentioned were independence, responsibility for learning, taking the correct information from text, exploring ideas and working as a group 50% of the teachers were positive about this style. 100% would be happy to incorporate these strategies into schemes of work.	100% felt that it had no additional effect on attainment and/or scientific understanding compared to their usual teaching style. 25% of teachers gave no positive comments or elaborated on the development of skills. 75% conjectured that of the pupils who they believed to dislike inquiry-based strategies, it was due to completing work for other pupils in a group and not having a written document of facts to learn. 100% do not think these strategies should be used solely.

References

Abrahams, I., & Millar, R. (2008). Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education*, *30*(14), 1945-1969.

Adey, P. (1999). The Science of Thinking, and Science for Thinking: A Description of Cognitive Acceleration through Science Education (CASE). Innodata Monographs 2.

Anderson, O. R., & Demetrius, O. J. (1993). A flow‐map method of representing cognitive structure based on respondents' narrative using science content. *Journal of Research in Science Teaching*, *30*(8), 953-969.

Arter, J. A., & Salmon, J. R. (1987). Assessing Higher Order Thinking Skills. A Consumer's Guide.

Association for Science Education (ASE). http://www.ase.org.uk/resources/science-upd8/

Bischoff, P. J., & Anderson, O. R. (2001). Development of knowledge frameworks and higher order cognitive operations among secondary school students who studied a unit on ecology. *Journal of Biological Education*, *35*(2), 81-88.

Brown, S. L., & Melear, C. T. (2006). Investigation of secondary science teachers' beliefs and practices after authentic inquiry‐based experiences. *Journal of Research in Science Teaching*, *43*(9), 938-962.

Burns, J. (2012, November 14). Exam success makes children happy, argues Michael Gove. *BBC News education reporter.* Retrieved from http://www.bbc.co.uk/news/education-20319008

Christodoulou, D. (2014). *Seven myths about education*. Routledge.

Crawford, B. A. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of research in science teaching*, *37*(9), 916-937.

Crawford, B. A., Krajcik, J. S., & Marx, R. W. (1999). Elements of a community of learners in a middle school science classroom. *Science Education*, *83*(6), 701-723.

Daud, N. M., & Husin, Z. (2004). Developing critical thinking skills in computer‐aided extended reading classes. *British Journal of Educational Technology*, *35*(4), 477-487.

DCSF (2008). Developing critical and creative thinking: in science. *The National Strategies*. Nottingham:DCSF

Dewey, J. (1899). *The school and society*. Chicago, Illinois: University of Chicago Press

DfE (2015). The link between absence and attainment at KS2 and KS4 2012/13 academic year. Crown Copyright.

DfE (2013). National curriculum in England: science programmes of study – key stage 3. Crown Copyright

Dillon, J. (2008). A Review of the Research on Practical Work in School Science. *King's College, London*, 1-9.

Ennis, R. H. (1962). A concept of critical thinking. *Harvard educational review*, 29, 128-136

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities.

Ennis, R. H. (1993). Critical thinking assessment. *Theory into practice*,*32*(3), 179-186.

Ennis, R. H. (2003). Critical Thinking Assessment. In D. Fasko (Ed.).*Critical Thinking and Reasoning: Current Research, Theory, and Practice* (pp.293-313). Cresskill, New Jersey: Hampton Press Inc.

Ernst*, J., & Monroe, M. (2004). The effects of environment-based education on students' critical thinking skills and disposition toward critical thinking. *Environmental Education Research*, *10*(4), 507-522.

Fasko, D. (2003). Critical Thinking: Origins, Historical Development, Future Directions. In D. Fasko (Ed.). *Critical Thinking and Reasoning: Current Research, Theory, and Practice* (pp.293-313). Cresskill, New Jersey: Hampton Press Inc.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching a meta-analysis. *Review of educational research*, *82*(3), 300-329.

Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B., Soloway, E., & Clay-Chambers, J. (2008). Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. *Journal of Research in Science Teaching*, *45*(8), 922-939.

Ginnis, P. (2001). *The teacher's toolkit: Raise classroom achievement with strategies for every learner*. Crown House Publishing.

Halpern, D. F. (2001). Assessing the effectiveness of critical thinking instruction. *The Journal of General Education*, *50*(4), 270-286.

Halpern, D.F. (2003). The 'how' and 'why' of critical thinking assessment. In D. Fasko (Ed.).*Critical Thinking and Reasoning : Current Research, Theory, and Practice* (pp.293-313). Cresskill, New Jersey: Hampton Press Inc.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

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

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99-107.

Hodson, D. (1998). Practical work in science: time for a reappraisal. In J. Wellington (Ed.), *Is this really what scientists do? Seeking a more authentic science in and beyond the school laboratory (pp 91-108).* London: Routledge

Horner, S (2011). Scientific inquiry: mere pedagogy or real science? And who decides? In Wellcome Trust (2011). *Perspectives on education: Inquiry-based learning*. London

Hughes, J.N. (1992). Review of the Cornell Critical Thinking Test. In J.J. Kramer, & J.C.Conoley, (Eds.), *The eleventh mental measurements yearbook* (pp. 241-243). Buros Institute of Mental Measurements. Lincoln: University of Nebraska

Keys, C. W., & Bryan, L. A. (2001). Co-constructing inquiry-based science with teachers: Essential research for lasting reform. *Journal of research in science teaching*, *38*(6), 631-645.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, *41*(2), 75-86.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction effects of direct instruction and discovery learning.*Psychological Science*, *15*(10), 661-667

Ku, K. Y. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking skills and creativity*, *4*(1), 70-76.

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction*, *18*(4), 495-523.

Lazonder, A. W., & Harmsen, R. (2016). Meta-Analysis of Inquiry-Based Learning Effects of Guidance. *Review of Educational Research*, 0034654315627366.

Liu, C. H., & Matthews, R. (2005). Vygotsky's Philosophy: Constructivism and Its Criticisms Examined. *International Education Journal*, *6*(3), 386-399.

Lunetta, V. N., Hofstein, A., & Clough, M. P. (2007). Learning and teaching in the school science laboratory: An analysis of research, theory, and practice.*Handbook of research on science education*, 393-441.

Lynch, S., Kuipers, J., Pyke, C., & Szesze, M. (2005). Examining the effects of a highly rated science curriculum unit on diverse students: Results from a planning grant. *Journal of Research in Science Teaching*, *42*(8), 912-946.

Malcolm, K.K. (1992). Review of the Cornell Critical Thinking Test. In J.J. Kramer, & J.C.Conoley, (Eds.), *The eleventh mental measurements yearbook* (pp. 243-244). Buros Institute of Mental Measurements. Lincoln: University of Nebraska

Mant, J., Wilson, H., & Coates, D. (2007). The effect of increasing conceptual challenge in primary science lessons on pupils' achievement and engagement.*International Journal of Science Education*, *29*(14), 1707-1719.

Marin, L. M., & Halpern, D. F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains.*Thinking Skills and Creativity*, *6*(1), 1-13.

McConney, A., Oliver, M. C., WOODS-McCONNEY, A. M. A. N. D. A., Schibeci, R., & Maor, D. (2014). Inquiry, Engagement, and Literacy in Science: A Retrospective, Cross-National Analysis Using PISA 2006. *Science Education*, *98*(6), 963-980.

Millar, R. (2004). The role of practical work in the teaching and learning of science. *High school science laboratories: Role and vision*.

Millar, R., & Osborne, J. (1998) *Beyond 2000: Science education for the future: A report with ten recommendations*. King's College London, School of Education.

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of research in science teaching*, *47*(4), 474-496.

Miri, B., David, B. C., & Uri, Z. (2007). Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking. *Research in science education*, *37*(4), 353-369.

Moseley, D., Elliott, J., Gregson, M., & Higgins, S. (2005). Thinking skills frameworks for use in education and training. *British educational research journal*, *31*(3), 367-390.

Murray, I., & Reiss, M. (2005). The student review of the science curriculum.*School Science Review*, *87*(318), 83-93.

NESTA – the National Endowment for Science, Technology and the Arts (2005). Real science: Encouraging experimentation and investigation in school science learning. London.

Ofsted (2013). Maintaining curiosity: A survey into science education in school. Crown Copyright.

Ofsted (2014). School inspection handbook. Crown copyright.

Oktay, J. S. (2012). *Grounded theory*. Oxford University Press.

Oxford Dictionaries. Retrieved from http://www.oxforddictionaries.com/definition/english/enquire accessed 27/06/14

Phillips, D. C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational researcher*, *24*(7), 5-12.

Pithers, R. T., & Soden, R. (2000). Critical thinking in education: A review.*Educational research*, *42*(3), 237-249.

Renaud, R. D., & Murray, H. G. (2008). A comparison of a subject-specific and a general measure of critical thinking. *Thinking Skills and Creativity*,*3*(2), 85-93.

Ryder, J. (2011). Scientific inquiry: learning about it and learning through it. In Wellcome Trust (2011). *Perspectives on education: Inquiry-based learning.* London

SCORE - Science Community Representing Education (Dec. 2008). Practical work in science: A report and proposal for a strategic framework. London.

Science and Technology Select Commitee10th report (Oct. 2006). Retrieved from http://www.publications.parliament.uk/pa/ld200506/ldselect/ldsctech/257/25706.htm#a12 accessed 06/06/14

Science and Technology Select Commitee10th report (Oct. 2008). Retrieved from

The National Committee of Inquiry into Higher Education (1996). Retrieved from http://www.leeds.ac.uk/educol/ncihe/natrep.htm accessed 07/06/14

Vygotsky, L. (1978). Interaction between learning and development. In M. Gauvain and M. Cole (Eds.) Readings on the development of children (1997). New York: WH Freeman and Company.

Walker, P. (2012, November 14). Tough exams and learning by rote are the keys to success, says Michael Gove. *The Guardian.* Retrieved from http://www.theguardian.com/politics/2012/nov/14/michael-gove-backs-learning-by-rote

Wallace, C. S., & Kang, N. H. (2004). An investigation of experienced secondary science teachers' beliefs about inquiry: An examination of competing belief sets. *Journal of research in science teaching*, *41*(9), 936-960.

Wellcome Trust (2011). *Perspectives on education: Inquiry-based learning*. London

Wellington, J. (1998). Practical work in science: time for a reappraisal. In J. Wellington (Ed.), *Practical work in school science (pp 3-15).* London: Routledge

Wellington, J., & Ireson, G. (2008). Science learning. *Science teaching*.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and instruction*, *16*(1), 3-118.

Willingham, D. T. (2009). *Why don't students like school?: A cognitive scientist answers questions about how the mind works and what it means for the classroom*. John Wiley & Sons.

Wilson, H., & Mant, J. (2011). What makes an exemplary teacher of science? The teachers' perspective. *School Science Review*, *93*(343), 115-119.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving*. *Journal of child psychology and psychiatry*, *17*(2), 89-100.

Wu, Y. T., & Tsai, C. C. (2005). Effects of constructivist-oriented instruction on elementary school students' cognitive structures. *Journal of Biological Education*, *39*(3), 113-119.

Yeomans, E. (2011). Inquiry-based learning – what is its role in an inspiring science education? In Wellcome Trust (2011). *Perspectives on education: Inquiry-based learning.* London