

**PHYSICS KEYWORDS: USING A QUASI-EXPERIMENT ALONGSIDE PUPIL
AND TEACHER SURVEYS TO EVALUATE PEDAGOGICAL CHANGE IN GCSE
PHYSICS**

by

MICHAEL ROBERT JACKSON, BSc (Hons) PGCE

A thesis submitted to the University of Birmingham for the degree of MASTER OF
ARTS IN RESEARCH (EDUCATION)

School of Education
College of Social Sciences
University of Birmingham
August 2014

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

The research addressed the question “Are there discernible benefits to pupils in physics when science pedagogy includes vocabulary-based activities?” This was achieved by using a quasi-experimental design in conjunction with a survey. 100 keywords from the EdExcel GCSE science course physics 1 module were targeted in physics lessons using a set of keyword sheets. Pupils’ ability to define 50 of these keywords was tested before and after teaching using a multiple-choice test. Interviews were then used for a sample of pupils and staff in order to develop understanding of the potential benefits.

Analysis of pupil scores showed a link between the teaching of the module and an increase in pupil scores on the keyword test. Standard deviation increased from 8.7 to 9.7 despite a 9.9 mean increase in scores. The lowest score in each set shows only an increase of one whilst the highest score is 48, two short of a perfect score (compared to 45 before the teaching). Multiple factors appeared to be affecting pupils’ ability to define words, including prior familiarity with the words. The interviews showed that pupils and staff felt that comprehension style activities had helped to develop pupil vocabulary though more extensive vocabulary activities were less useful. Recommendations were made for future development of studies concerning vocabulary interventions, including the use of science glossaries.

Acknowledgements

This thesis owes particular recognition of a number of people, especially due to the study being carried out alongside my role as a full-time Head of Science.

Primarily my thanks go to Dr. Tonie Stolberg of the University of Birmingham for his role as my research supervisor. Dr Stolberg's advice and calm reassurance have made not only this thesis possible but have helped me to balance the demands of academic study and my role as a Head of Department in school throughout the year and also ensured that I remain full of enthusiasm for academic study and my research.

My thanks also go to Dr. Paul Warmington, my research supervisor for my first year, who guided me towards my research focus. Also thanks to Prof. Ann-Marie Bathmaker and Prof. Peter Davies for their support, advice and enthusiasm whilst narrowing my research focus.

Also thanks to all of the staff at the University of Birmingham (and particularly the School of Education) who have supported this research, the lecturers who gave of their time on Saturdays to provide research lectures and the staff of the School of Education library.

Thanks also go to the Head teacher of the school in which this research took place, for his support in allowing me to proceed. Also thanks to the staff in my faculty who have supported this research on both a personal and professional level. My thanks to the pupils who took part in the focus groups and gave willingly of their time, with their parents' support. Their honesty and maturity were invaluable.

Finally but by no means least my thanks to those who have lent their personal support. My thanks go to Kerry for her interest in my work and her support as I have completed it. Thanks also go to my family for their ongoing support on this path. Also thanks to Natalie for ensuring I had the encouragement and belief to apply to start this study two years ago.

Many others are involved in the culmination of this work – the teachers who taught me from Reception through to Sixth Form and my undergraduate tutors. I am lucky to have had an overwhelmingly positive time in education for virtually every day of my life and I hope that this work can add something small to the enjoyment and development of pupils in Science and beyond!

Table of Contents

Chapter 1 – Introduction	1
1.1 Rationale	1
1.2 Research objective and questions	8
Chapter 2 – Literature Review	10
2.1 Introduction	10
2.2 Identifying useful directions of research and a gap in the literature	11
2.2.1 The relative importance of literacy in science learning	11
2.2.2 The importance of pupil motivation, aspiration and engagement	16
2.3 The role of vocabulary in teaching in science lessons	20
2.3.1 Changing vocabulary demands in science	20
2.3.2 Understanding the demands of vocabulary	21
2.3.3 Previous experimental research on vocabulary	26
2.4 Developing language in the classroom	31
2.4.1 Approaches by other researchers	31
2.4.2 The “Four Strands” approach to vocabulary development	35
2.4.3 Vocabulary development activities	40
2.5 Testing knowledge of vocabulary	41
2.6 Conclusions from the review of literature	43
Chapter 3 – Research Design and Methodology	46
3.1 Introduction	46
3.2 Selection of research design frames	46
3.3 Development of research design frames	51

3.4 Quasi-experimental design	51
3.4.1 Description of the quasi-experimental design frame	51
3.4.2 Ethical considerations	54
3.4.3 Population and sampling	54
3.4.4 Group size	56
3.5 Survey design frame	57
3.5.1 Description of the survey design frame	57
3.5.2 Population and sampling	59
3.5.3 Group size and composition	60
3.5.4 Ethical considerations for the survey	61
3.5.4.1 Ethical considerations for children as research participants	63
3.5.4.2 Ethical considerations for teachers as research participants	65
3.5.4.3 Ethical considerations of access and conflicts of interest	65
3.6 Order of research	66
3.7 Research methods	68
3.7.1 Changes to pedagogy and the intervention tool	68
3.7.1.1 Selection of keywords for each worksheet	68
3.7.1.2 Worksheet format	69
3.7.1.3 Use of the worksheets in lessons	69
3.7.1.4 Staff CPD for use of the worksheet	70
3.7.2 Quasi-experiment data collection	71
3.7.2.1 Testing understanding of the keywords	71
3.7.2.2 Recording of test data	75
3.7.3 Processing of data from the quasi-experiment	75

3.8 Survey data collection	77
3.8.1 Interview methodology	78
3.8.2 Nature of the interview and schedule of questions	81
3.8.3 The role of the interviewer	83
3.8.4 Considerations surrounding interviewees	83
3.8.5 Validity	84
3.8.6 Piloting of the focus groups	85
3.8.7 Analysis of the data	85
3.8.7.1 Introduction to the analysis of focus groups	85
3.8.7.2 Supplementary information from the focus groups	86
3.8.7.3 Analysis of the “unspoken” aspects of the focus groups	87
3.8.7.4 Using transcripts to address the research questions	89
3.8.7.5 Thematic analysis of the interviews	90
Chapter 4 – Analysis of Research Question 1	92
4.1 Introduction to the analysis of findings	92
4.2 Analysis of data for all pupils by pupil score	92
4.3 Analysis of data by word score	99
4.3.1 Improvement in score for individual words	100
4.3.2 Analysis of pre-test scores for individual words	108
4.4 Familiarity of vocabulary	111
4.5 Pupils with English as an additional language (EAL)	117
4.6 Effects of the multiple-choice test	118
4.7 Teacher-effect	122
4.8 Concluding points regarding the first research question	123

Chapter 5 – Analysis of Research Questions 2 and 3	126
5.1 Introduction	126
5.2 What worthwhile benefits do pupils perceive from the change in pedagogy?	126
5.2.1 Pupils perceptions of the keyword sheets	126
5.2.2 Pupil perceptions of the structure and use of the keyword sheets	129
5.2.3 Pupil perceptions of testing associated with the intervention	131
5.2.4 Pupil motivation towards keyword sheets	131
5.2.5 Conclusions in response to the second research question	133
5.3 What worthwhile benefits do teachers of the pupils perceive from the change to pedagogy?	135
5.3.1 Teachers' perceptions of the use of the keyword sheets	135
5.3.2 Staff perceptions of the keyword sheet layout	137
5.3.3 Teachers' perceptions of testing in connection to the keyword sheets	137
5.3.4 Teachers' perceptions of pupil motivation	138
5.3.5 Final conclusions in response to the third research question	139
Chapter 6 – Further Discussion	142
6.1 Lesson content	142
6.2 Pupil motivation	143
6.3 Revision	145
6.4 Practical work	147

6.5 Mathematics and STEM in science	149
6.6 Lesson scheduling	151
6.7 Concluding points regarding further discussion	151
Chapter 7 – Final Conclusions and Future Directions	153
7.1 Introduction	153
7.2 Are there discernible benefits to pupils in physics when science pedagogy includes vocabulary-based activities?	153
7.2.1 Effect size	154
7.2.2 Cost of interventions	155
7.3 Future research directions	156
7.3.1 Development of the intervention	156
7.3.2 Development of the word test	157
7.3.3 Development of the survey interviews	158
7.4 Summary	159
List of references	160
Appendices	174
Appendix A – Physics 1 keyword list	
Appendix B – Words focussed upon in the physics intervention	
Appendix C – Ethics sheets	
Appendix D – An example of a keyword sheet	
Appendix E – Word Test	

List of Illustrations

Figure 3.1 – Order of research	66
Figure 4.1 – A scatter-graph showing the correlation between pre-test and post-test Scores for each Pupil	96
Figure 4.2 – A frequency histogram for pupil score pre-test and post-test	98
Figure 4.3 – A comparison of pre-test word scores to improvement in score	104

List of Tables

Table 3.1 – Research questions and principal design frames	50
Table 3.2 – The experimental variables	52
Table 3.3 – Pupil teaching groups	57
Table 4.1 – Descriptive statistics for pupil scores pre-test and post-test	93
Table 4.2 – Descriptive statistics for success in terms of individual words	100
Table 4.3 – Word scores ordered by percentage of pupils defining each word correctly	103
Table 4.4 – Pre-test word scores ordered by the percentage of pupils defining each word correctly	109
Table 4.5 – Keywords grouped by possible previous pupil encounters	113
Table 4.6 – An arbitrary scale for familiarity of science keywords in GCSE Physics	115
Table 4.7 – Multiple-choice answers selected for each keyword	119

List of Definitions and Abbreviations

AABs – Attitudes, Aspirations and Behaviour

ADHD – Attention Deficity Hyperactivity Disorder

AEAL – Advanced English as an Additional Language

BERA – British Education Research Association

CPD – Continuing Professional Development

DfE – Department for Education (UK Government)

EAL – English as an Additional Language

ELL – English Language Learner (in the USA)

EM – Electromagnetic

GCSE – General Certificate of Secondary Education

OECD – Organisation for Economic Co-operation and Development

OFSTED – The Office for Standards in Educations, Children's Services and Skills

PGCE – Post Graduate Certificate in Education

SD – Standard Deviation

STEM – Science, Technology, Engineering and Maths

Chapter 1 – Introduction

This research arose from my observations and reflections whilst working in a secondary school. This introduction outlines this personal perspective so as to provide both a rationale for the work and to allow the reader to evaluate the author and researcher's own perspective prior to the study.

1.1 Rationale

As a Head of Faculty for science in a city comprehensive school for girls in Birmingham, England, I had become increasingly aware of the demands that the science subjects placed on pupils at GCSE level. In response to this the department had been developing the use of an approach based on Bloom's (1956) taxonomy to ensure that pupils were being taught at an appropriately challenging level. This approach, as well as encouraging pupils to work to a higher level, had made the need to develop underlying knowledge of terminology (Bloom, 1956, p.201) more apparent as well. This meant that the department had begun to focus on developing scientific vocabulary in order to try and provide pupils with a stronger grounding upon which to build towards higher levels of thinking in science. Whilst developing scientific ability was a worthy goal in itself, in a secondary school the hope was also to improve examination results at GCSE level.

The media had also been focussing on the link between reading and examinations at this time. Burns (2013) reported for BBC News that in 2012 13% of seven-year-olds "were not at the expected reading level." She also reported that "almost a quarter

(24%) of children eligible for free meals did not reach their expected level, compared with only 10% of children from better-off families.” This brought together the author’s concerns about the link between reading, progress and socio-economic background starting from a very young age. Burns source was the Save the Children report entitled “Too Young to Fail” which claimed in more detail (Save the Children, 2013, p.5) that there was an inverse link between family income and achievement at GCSE, having used data from the National Pupil Database. This was particularly of concern in a school with many families of low income; a challenge facing schools from across this country as well as abroad.

New GCSE courses were introduced in 2011, leading to a particular focus on these courses and the new challenges for the pupils (and staff teaching them). Science GCSE textbooks, such as Pearson’s *EdExcel GCSE Science* (Levesley, 2011) showed a significant use of complex scientific language – around 200 words were printed in bold (identifying them as keywords) for each module. This demonstrated the level of challenge involved in ensuring that pupils were comfortable with the technical vocabulary in order to meet the higher level challenges and ultimately gain the higher grades. However, it was difficult to assess the relative complexity of language for pupils. Pupils came to the school from numerous local Primary schools whilst others sometimes moved to the area from elsewhere in the country and overseas so there was considerable variation in the prior experience of language in students’ prior learning. The school, in which this study is set, had not specifically attempted to measure complexity of vocabulary in science or the ability of pupils to understand the required keywords, other than through incidental encounters during the teaching of science courses.

As recently as 2010 pupils could sit GCSE examinations in Core Science by completing multiple-choice question papers. The new specifications introduced for 2011 required pupils to answer a variety of question styles, including 6-mark questions which involved writing short essays that were assessed against a three-tier criteria that includes spelling, punctuation, grammar and clarity of expression. An example of a higher level question and the higher level marking criteria for a physics question about the electromagnetic spectrum in the 2013 summer examination series is shown below:

Question

***Radiation from different regions of the electromagnetic spectrum can affect the human body in many ways.**

Discuss the different ways in which excessive exposure to electromagnetic radiations of various frequencies may cause damage to the human body.

(Pearson Education Ltd, 2013a)

Mark scheme

- A detailed description e.g. gives most of the correct radiations with links to detail of the damage AND explains the link between frequency and energy/danger, e.g. microwaves heat up the water in cells; UV can cause

cataracts; gamma rays are the most penetrating and can mutate cells inside the body because they have the highest frequency;

- The answer communicates ideas clearly and coherently uses a range of scientific terminology accurately;
- Spelling, punctuation and grammar are used with few errors.

(Pearson Education Ltd, 2013b)

There was a need to use scientific terminology accurately, which would require the pupil to have a sound understanding of the vocabulary used in the question as well as that which was required for the answer. Examples here included “radiation”, “regions”, “electromagnetic spectrum”, “excessive”, “exposure”, “frequency”, “microwaves”, “ultraviolet”, “cataracts”, “gamma rays”, “penetrating”, and “mutate”. The vocabulary requirements increasingly became apparent from the text books, specifications and past papers over time but the level of challenge this language created for the pupils was harder to measure. Perera (1980) discussed the judgement of complexity of language in detail. She (1980, p.152) listed a subjective teacher assessment as a suitable method for doing this. Although there are more mathematical methods, Perera (1980, p.154) warned that “they are unable to take account of conceptual difficulty and interest level.” As such at this early stage the judgement on word complexity was subjective but clearly worthy of further consideration, potentially quantitatively so as to accurately assess pupils’ needs.

A further change to examinations was the move to terminal assessment after many years of a modular approach, with re-sit possibilities throughout the two years. This

meant that as of the 2013/14 academic year pupils were required to sit all of their examinations for each course in the summer of the year that they completed the course. In science subjects this meant three examinations in Year 10 and three examinations in Year 11, or nine examinations in Year 11 for those sitting the separate science examinations over the two years. The need to retain knowledge for a longer period of time suggested that the impact of a pupil's secure word knowledge on final grades could be higher. However, this concern was embedded in the complex factors that affect pupils' performances in examinations, such as their own revision and performance under higher pressure, with no opportunity to learn from failure.

In addition to these concerns was the move towards a wider ability range of pupils studying these heavily examined GCSE courses. The current Government made clear its belief that the majority of pupils are suited to a standardised approach at GCSE level. In a 2013 report the Department for Education (DfE, 2013b, p.12) said "Our modelling suggests that around 1.2% of pupils will not be recognised in the Progress 8 measure because their particular needs mean that they cannot enter any GCSEs or high value vocational qualifications." In reality this meant that league table measures for schools from 2016 onwards would not include vocational alternatives to be counted for science. This meant a change from pupils who would not be likely to achieve a C grade at GCSE being entered for the vocational, assignment based, BTEC Applied Science (or an OCR National) to almost all pupils studying the more traditional GCSE courses. As such there was a greater need to understand pupil interactions with scientific vocabulary and to look for ways to help pupils to acquire it, irrespective of their perceived ability level or prior achievements.

Brown (1987, p.3), discussed the role of vocational courses and said of the introduction of such courses in the Thatcher era, "...it is motivated more by an attempt to maintain (indeed extend) educational and social inequalities than to equip pupils for adult life." This was seen to ethically support the aforementioned switch from vocational courses to more traditional GCSE courses and further heighten the need to find solutions to the problems, such as vocabulary, that arise rather than simply looking to alternative courses.

The discussion of league table impacts and Government policy could have led to a wider debate of the ethics of decisions about courses and the surrounding teaching. However, this study needed to reflect the reality that schools are judged heavily by league tables and are required to respond to them. As such the debate is largely irrelevant and there was a need to focus on the challenge rather than be caught up in debating whether something beyond teachers' control is right or wrong. The DfE vision (DFE, 2013a) in 2013 was for "a highly educated society in which opportunity is more equal for children and young people no matter what their background or family circumstance." It seemed reasonable to conclude that despite the politics and results-driven culture of modern education in the United Kingdom there was still a shared aim between education professionals of striving to provide the best education for all pupils. It also seemed that a focus on the academic courses would be agreeable to those from disparate viewpoints on the vocational education debate.

Despite this shared aim, which had been consistent across many decades, Brown (1987, p.3) had said that "ordinary kids have been largely ignored and misunderstood...because [of] their 'invisibility' and apparent conformity...associated

with female pupils.” Working in a girls’ school made this poignant. The IOP (2012, p.5-7) noted their concern with the uptake of physics by girls from comprehensive schools, though they do note that girls from a single-sex school were “two-and-a-half times more likely to go on to do A-Level physics if they came from a girls’ school rather than a co-ed school...” This conception of poor uptake of physics by girls must be regarded with caution however as Mujtaba and Reiss (2013) found that teacher expectation and encouragement are major factors in girls’ progression in science, as opposed to any link to language or the inherent nature of the subject (when compared to boys). Archer *et al.* (2013) also found that younger girls studied (from ten to fourteen years old) did not find science aspirations in keeping with their own constructions of femininity or their personal sense of themselves as learners. This appeared to be particularly the case for working class pupils. These pieces of research began to highlight the complex interactions that happen within the subject and the variables that make analysis in this sphere of research complex. This added weight to studying gender discretely in science whilst being cautious of preconceived ideas surrounding girls.

Further analysis of the impact of vocabulary on pupils led to the work of Raz and Bryant (1990) who found that social class had a significant impact on phonological awareness and reading ability at school-age. Whilst not directly linked to vocabulary this furthered the idea of social class as a factor which might influence pupils’ knowledge of language in each subject. In considering the impacts of a working class setting, Brown (1987, p.11) said “Most of the writings on education and the working class have been presented within a wider debate about social justice, inequality and equality of opportunity.” The intent here was not to focus on the wider

debate but upon the narrower impact in science and particularly upon vocabulary, despite the potential relevance of the wider issues. Gorard, See and Davies (2011, p.7) said that “Understanding the reasons for the poverty ‘gradient’ and devising approaches and suggested behaviour changes that help reduce it are therefore directly relevant to current policy and practise.” This did not provide insight into the effects of the social setting of the school but does encourage the search for approaches to overcoming the arising issues, of which vocabulary could have been one.

In summary, this research aimed to improve attainment of pupils, regardless of their socio-economic background, using further insight into the relative importance of language with direct relevance to the current examinations. There appeared to be a gap in research of this nature and that similar research which did exist would require re-consideration if it were not carried out in the run-up to the examinations in the summer of 2013 due to introduction of examination questions which required more precise and extensive use of language. The potential here was to develop a vocabulary-based intervention and to investigate whether such an approach would help pupils studying GCSE level science. Whilst studied in one school it would then have application to other schools teaching at a similar level, particularly those offering GCSE physics courses.

1.2 Research objective and questions

This objective of the work was to further develop the understanding of the impact of language in science upon pupils’ progress in the subject. This was achieved by

researching the impact of developed teaching pedagogy both quantitatively and qualitatively. The primary research question was:

Are there discernible benefits to pupils in physics when science pedagogy includes vocabulary-based activities?

This was answered via the following sub-questions:

1. Do pupils show an improved ability to define keywords following the introduction of a vocabulary-based intervention?
2. What worthwhile benefits do pupils perceive from the change in pedagogy?
3. What worthwhile benefits do teachers of the pupils perceive from the change to pedagogy?

Note that the choice of physics as the selected module was due to a variety of factors, including observed difficulties in particular with this unit alongside the convenient organisation of the teaching of this unit which fitted the research schedule given its timing in the academic year. At this time classes were established with their teachers and all classes studied the module at a similar time. The aforementioned discussion of girls and physics also provided additional insights but this was not a driving factor in the decision to choose this unit. The keywords chosen were based upon the GCSE textbook (Levesley, Ed., 2011) by the exam board's (EdExcel) parent company Pearson as discussed in the research design and methodology in Chapter 3, after consideration of the literature relevant to these questions in Chapter 2. The later chapters will provide analysis of the findings in relation to the research questions and the wider findings.

Chapter 2 - Literature Review

2.1 Introduction

The literature review begins by further exploring the motivations and aspirations of pupils in science which added understanding to the literature gap concerning scientific vocabulary and the need for further research that had been suggested in the introduction. This led to studies which had already been conducted and had shown some success through direct vocabulary interventions. The wider sphere of language skills was then considered in order to develop an appropriate intervention tool. This in turn led to consideration of how to develop a test for vocabulary knowledge and then literature which further aided the development of methodology. Some of the literature was not science specific as useful research was found concerning impacts upon English language skills that could be applied to science teaching. Research relating to other subjects and EAL (English as an Additional Language) learning was also considered where appropriate links to science vocabulary were identified and there was potential for application to this study.

Given that development of vocabulary can be influenced by a number of factors beyond the classroom, literature was also considered where it had provided insight into the effect of those factors on pupils' vocabulary and wider learning in science. Wider work on pupil attainment and progress was also considered, though only that research which helped to develop an understanding of the results from this research and its potential impact on pupils.

Consideration was given to publications from the inspectorate of schools, OFSTED, and the Government. OFSTED in particular had carried out recent research across English schools, including subject surveys in science. This made their work of particular interest since it had covered similar school settings and the science subjects. Such research might be considered less reliable than academic papers and books due to potential political bias but the breadth of the work and its relevance mean that it was still been useful to reflect upon in this review.

2.2 Identifying useful directions of research and a gap in the literature

Whilst the background to this study had identified vocabulary as a central issue it was deemed important to research key current issues within the topic of teaching and learning in science in order to establish the importance of vocabulary as a point of research relative to other issues and also to ensure that vocabulary development took account of other key issues that might have an effect.

2.2.1 The relative importance of literacy in science learning

Francis, Hutchings and Read (2004) produced a report which was research commissioned by the Association of Maintained Girls' Schools. This report provided more detailed insight into success in science and in similar contexts to the school which would ultimately host this research. They (2004, p.ii) listed the priorities they found from schools "more successful" in science; a focus on teaching and learning, class discussion, metacognitive approaches, high expectations, targets and monitoring and continuing professional development. They also said that schools

“less successful” in science emphasized variety, factors relating to gender and practical work and examples. They noted “an overwhelming tendency for girls to see the relevance of science in their own lives in order to be engaged by it,” along with engagement from group work and practical experiments. This was interesting since they noted practical work to be important but also that it is an emphasis in schools that were deemed less successful. This aimed the intervention direction further towards “teaching and learning” and the aforementioned other factors which were considered a part of this wide theme. Francis, Hutchings and Read (2004) further noted that girls disliked “copying from the board or textbooks” as well as “Abstract, technical and mathematical elements” (2004, p.iii). They warned (2004, p.iv) that the science curriculum was too content heavy and called for review at Key Stage 4. Francis, Hutching and Read (2004) used questionnaires and focus groups to provide various information and conclusions about elements of the teaching of science, exemplifying the use of qualitative research to better inform teaching. They did not appear to consider vocabulary in science as an area of focus beyond discussion about the difficulty and technicality of science – which seemed to fit within the area of pupil dislikes. This began to highlight the research gap in the literature and demonstrated it in girls’ maintained schools in England. There seemed to be issues surrounding more technical scientific work but pupils did enjoy science where they were motivated and saw a relevance to the science being learnt or simply the important of being successful in science. The report did also highlight again the other numerous factors affecting pupils’ experiences of science.

The schools inspectorate in England, OFSTED, continued the theme of technical science work in their subject report for science. They recently developed updated

teaching categories and under the category of “Outstanding” (the top category) they said such a lesson would show that (OFSTED, 2014, p.3) “There is excellent practice which ensures that all pupils maintain high levels of literacy in science appropriate to their age.” Conversely, OFSTED (2014, p.3-4) also listed seven statements that would result in an “inadequate” judgement being made. This included “pupils’ progress in literacy.” Clearly OFSTED felt that literacy in science was of vital importance to effective teaching. OFSTED also said, under quality of teaching in science, that (OFSTED, 2014, p.7) “Pupils’ active participation in their learning secures outstanding progress across all aspects of the subject because teachers use a very wide range of innovative and imaginative resources including local contexts, along with well-chosen teaching strategies.” In other words OFSTED believed that it was necessary to use literacy alongside a wide range of resources in an active way to secure strong pupil engagement to secure progress in science. They applied the converse to their inadequate category saying (OFTSED, 2014, p.8) “Teaching fails to engage pupils’ interest in science. The content of science lessons is not often in contexts that relate to pupils’ lives and the relevance of science is not made apparent.” Again under quality of curriculum they said (OFSTED, 2014, p.9) “The contexts in which science is taught are relevant to pupils’ lives, capture their interest...” and “There are productive links with other subjects...including...English.” Seemingly OFSTED agreed with the findings in other literature that literacy was an area of importance and should be embedded within engaging lessons.

The *OFSTED 2012/13 Annual Report* (2013, p.2) reflects inspections of 7,905 of the 21,957 schools in the country and 1,334 of the 3,328 secondary schools. They (OFSTED, 2013, p.6) noted that “English and mathematics are not taught well

enough...” Whilst this was aimed at those subjects specifically, they (OFSTED, 2013, p.12) went on to support their comments from the science survey report (OFSTED, 2014) by saying inspectors and schools should expect that “pupils’ responses demonstrate sufficient gains in their knowledge, skills and understanding, including of literacy and mathematics.” This further evidenced the need to develop literacy further in many schools across England, furthering the value of this study.

Wilms (2003, p.34) discussed literacy further, concluding that “Over one-half of all students are engaged at school and have strong or at least average literacy skills.” He described the other half as either having fairly high literacy skills but low sense of belonging (20%), very poor literacy skills (10%) or regularly absent from school with weak literacy skills (10%). In other words Wilms deemed about 20% of pupils to have issue with literacy skills. Whether there was a causal factor in absence or vice-versa is unclear. Nonetheless these groups required appropriate attempts to improve their science (and other) learning and weak literacy skills were identified as an area of weakness.

Wilms (2003, p.48) also acknowledged that “Students from low socio-economic families are more likely to be disaffected from school, as are students who attend schools that have a high percentage of students of low socio-economic status. As these risk factors compound, students from low socioeconomic status families are even more likely to be disaffected from school.” This further evidenced the need to consider the socio-economic groups of the pupils and tackle the issues they face. Within the classroom it had often been observed that some pupils had less general knowledge and experience of the wider world, having difficulties with understanding

of objects such as pylons or features of the landscape such as quarries. This continued to focus the research towards use of precise vocabulary that might be less well developed for some pupils.

Pampaka *et al.* (2012, p.473) compiled a study of mathematics which although it was based in a different subject was useful for insights into physics in particular, given the level of use of mathematics in the subject. They found a need to “address the current concerns about teaching-to-the-test and its association with declining dispositions towards further study of mathematics and the consequences for choice of STEM subjects at university.” There was concern for science within the context of STEM (Science, Technology, Engineering and Mathematics). They went on to say (Pampaka *et al.*, 2012, p.474) “...it is widely believed that the drive to raise standards can be counterproductive for dispositions, especially when it has the effect of narrowing teaching practices. But hard evidence of the way that teaching mediates the effect of policy on different learning outcomes such as dispositions is largely missing from the literature: a gap we hope to inform.” There were several important elements to this discussion, including being wary of resorting to teaching methods which might suit the examinations but lead to a lowering of enthusiasm for science. Of relevance to the vocabulary and literacy link, they suggested methods that included connecting mathematical and scientific knowledge and connecting mathematical understanding to “dialogic and discussion-based communicative” mathematics (Pampaka *et al.* 2012, p.474).

2.2.2 The importance of pupil motivation, aspiration and engagement

Gorard, See and Davies (2011) provided a key reference point in the search for interventions in teaching and learning with quantitative analysis and review of 166,000 research reports aimed at identifying the impacts of attitudes and aspirations. Teachers had often remarked about the effect of attitude ahead of examinations and the role of aspirations so this seemed a logical point to consider in the search for successful impact tools and gaps in the research. Gorard, See and Davies (2011) provided helpful pointers towards the areas of attitudes and aspirations that appeared to show significant correlations across this multitude of studies from various countries, suggesting that any response to their findings would apply to some extent internationally. For instance, they (2011, p.116) found that individual motivation had a medium, positive effect on attainment and that individual attitude had only a weak, mixed association with attainment. Furthermore there was no evidence of intervention success. This was immediately useful in suggesting that motivation was not as critical as it was sometimes perceived by teachers. It also suggested that any intervention would be worthy of trial across pupils of varying motivation. Motivation was not however to be confused with engagement, which Francis, Hutchings and Read (2004) found to be important in girls' schools in particular. Gorard, See and Davies' work in combination with the work of Francis, Hutchings and Read (2004) showed that the effects of interventions surrounding motivation were not clear overall, with motivation to reach a goal more important than actual subject enjoyment. This suggested that attempting to use motivation as an intervention would be complex and require an in-depth understanding of the effect on pupils, one which would suit a qualitative analysis (at least in part). It did seem clear

that attitude ought not to be a primary concern and that pupils could benefit from interventions irrespective of their attitudes.

Galton, Gray and Ruddock (2003, p.56) made a further relevant point, saying that “...the pupils who performed best often had the poorest attitudes to science.

Motivation levels of these pupils were high – they wanted to do well – but only to prove they could pass the tests and do well in future examinations – not because they liked science as a subject.” This further acted as a deterrent to overemphasising the importance of attempting developing enjoyment and motivation directly. It suggested achievement was linked to other factors. However, this study would continue to note the potential importance of those things for general well-being and as a non-attainment based objectives which are not studied further within the confines of this research but could be further analysed if continuing to evaluate the impact of interventions more widely.

Another wide-ranging study was produced for the Rowntree Foundation by Goodman and Gregg (Eds., 2010). They concluded that there were three major areas in which policy might contribute to reducing educational inequalities and listed those as:

- a) reducing children’s behavioural problems; improving coping and management capabilities for risky behaviours, conduct disorder and attention deficit hyperactivity disorder (ADHD);
- b) helping children from poorer families to believe that their own actions and efforts can lead to higher educational outcomes;

c) raising children's aspirations and desire for advanced education – from primary schooling onwards...”

(Goodman and Gregg, Eds., 2010, p.8)

Goodman and Gregg's (Eds., 2010) work suggested that research efforts would be better focussed on self-belief and raising of aspirations – something which to an extent Gorard, See and Davies (2011) and Francis Hutchings and Read (2004) appeared to conclude as well. However they did not comment on the effect of language and vocabulary directly. Nonetheless their research did suggest that interventions might benefit from considering how children can be brought to believe that their own actions could lead to higher outcomes. Utilising vocabulary development seemed apt since pupils could be encouraged to use this as the aforementioned grounding for work at a higher level. These findings also warned of factors affecting educational inequality that might not be tackled by a vocabulary intervention such as low aspirations and poor self-belief that became apparent in some pupils through comments in focus groups and attitudes towards the keyword test.

In his OECD (The Organisation for Economic Co-operation and Development) report, Wilms (2003, p.29-30) did discuss literacy, suggesting it was not strongly linked to school engagement and participation. However there was potential for the investigation of a weak link to be studied further as Wilms (2003, p.33) said “Students’ participation is weakly related to measures of literacy performance.” This did not preclude the possibility that literacy might have been affecting attainment in science even if participation and engagement would not be significantly altered. It

also suggested that participation could have been affecting literacy performance. Thus it was conceivable that pupils with lower participation would benefit from attempts to improve their literacy and vocabulary might play a role in this. Furthermore a literacy-based intervention would need consideration to be given to engagement, perhaps through other elements of the lesson. This again connected back to similar ideas from Gorard et al. (2011) and Francis, Hutchings and Read (2004).

Dillon (2011, p.5) discussed this idea of engagement in depth and pointed out that “Even in an age when the plurality of TV channels and internet websites seems endless, blockbuster series featuring science, particularly nature or space, can generate huge audiences.” This raised the idea that complex language does not necessarily detract from interest – though clearly direct comparison between film and education had not been fully considered; films have been popular in other languages or without language at all, suggesting that language might not always be critical in enjoyment of a film. Nevertheless language was not a barrier to interest necessarily. There were clearly complex links between complexity of language and attitudes towards that thing which utilised the language. Gorard, See and Davies (2011) to an extent also suggested that interest can overcome these complexities. This further highlighted the desire for an intervention to work alongside ways of keeping science interesting and engaging even if the direct impact on results had not been shown.

At this stage the literature was evidencing a need to ensure pupils were engaged and motivated. It was also becoming clear that language skills were factors associated with working-class learners. Furthermore effective teaching and learning

were found in successful science departments. The importance of word groups and use of language in STEM subjects was beginning to become apparent and so further research into the use of language and vocabulary would be a useful next step in helping pupils develop in science through intervention in current teaching.

2.3 The role of vocabulary in teaching in science lessons

2.3.1 Changing vocabulary demands in science

The complexity of science vocabulary has been studied and discussed at length in various texts as well as the introduction to this research. Its importance was brought to prominence by the Department for Education's (DfE) reform of GCSEs. In June the DfE (2013c) published its consultation on science GCSE subject content and assessment objectives. The very first section on Biology stated:

Prokaryotic and eukaryotic cells

- ☐ describe a cell as the basic structural unit of all organisms
- ☐ describe the main sub-cellular structures of eukaryotic cells (plants and animals) and prokaryotic cells
- ☐ relate sub-cellular structures to their functions, especially the nucleus/genetic material, plasmids, endoplasmic reticulum, mitochondria, ribosomes, chloroplasts and cell membranes
- ☐ **evaluate the impact of electron microscopy on our understanding of sub-cellular structures including the nucleus, plasmids, mitochondria, chloroplasts and ribosomes**
- ☐ explain the aseptic techniques used in culturing microorganisms.

(DfE, 2013c, p.5)

This was a small sample of the approximate nine subsequent pages of similar language. Clearly the language used and required by pupils would be very complex.

This suggested a need to find ways to support pupils over to this obstacle. This contrasted with the much less specific 2006 guidance:

a) Cells and growth

Chemical reactions essential for life and growth take place inside cells.

Differences between plant

and animal cells lead to different patterns of growth and development.

(QCA, 2006, online)

This demonstrated the increased level specified by the Government and its agencies responsible for science education in England but did not clarify the demands this placed upon pupils.

2.3.2 Understanding the demands of vocabulary

Wellington (2000b, p.167) posed a question of pertinence here: “How many pupils, confronted by a science textbook or by a blackboard covered in scientific prose, are as confused as Alice when she first read ‘Jabberwocky’?” Wellington was of course referring to Alice in Wonderland and making the point that science can often sound very confusing to those with less developed knowledge of scientific language. He further warned (Wellington, 2000b, p.167) that “Unfortunately, many concept words in science do not, and cannot, acquire meaning as easily as a word like ‘trachea’”. Wellington’s statement emphasised the impact of the type of word that is being studied in science upon pupils’ ability to acquire meaning for the word. However later findings provided a complex picture of which of those words had been found more or less difficult for pupils to acquire. The word type was not the only factor found to have been at work. These findings were supported by Ross, Lakin and

Callaghan (2000, p.25) who in concurring said that “When children use words they often mean something completely different from the accepted scientific meaning.”

They went on to say that (Ross, Lakin and Callaghan, 2000, p.27) “Our task, as science teachers, is to become aware of these built-in but naïve ways of thinking...”

This made clear not only the complexity of words but the need to be aware of the meanings that pupils attached to them. Braund (2008, p.5-6) supported this idea in a more applied sense, saying that “Continuity...should require teachers receiving pupils from another class or key stage to give attention to the learning the children have already received...” This again supported the idea of being aware of prior conceptions. This would lead to vocabulary activities enabling teachers to explore prior understanding whilst developing the appropriate meaning for science lessons.

Bloom’s (1956, p.29) work on a taxonomy of learning domain was useful as it took the complexity of language a stage further in saying “The emphasis on knowledge as involving little more than remembering or recall distinguishes it from those conceptions of knowledge which involve “understanding,” “insight” or which are phrased as “really know,” or “true knowledge.” Here Bloom identified the difference between recalling meaning of words and the ability to understand and use them. He placed knowledge or terminology as one of the least demanding skills in his taxonomy (Bloom, 1956). According to Bloom, more increasingly demanding skills were (in order of increasing difficulty) comprehension, application, analysis, synthesis and evaluation. This extended the idea of different words having different levels of difficulty for pupils and identified different uses of those words as having different skill demands. The skills demanded in lessons and in examinations were clearly varied such that recall was an appropriate starting point when tackling

vocabulary in terms of skills level but was not the end point in developing use of those words. Bloom (1956, p.36-7) explained that curriculum development required decisions about the extent of use of the language according to this taxonomy. This was clearly another consideration to account for in planning an approach to tackle vocabulary difficulties in science. The approach would also need to take account of the restrictions caused by curriculum and examination specification demands that would be compounded by issues such as the time available to cover the content and the speed at which pupils acquired and were able to use the vocabulary. Bloom (1956, p.36-7) expanded on this suggesting that organisation of learning might suit the organisation or specialist (school or teacher perhaps) or the learner's stage of development. The challenge then would be to organise the learning of the "knowledge of terminology" and to allow the teachers to attempt to manage the limitations of time alongside the needs of the learner in terms of developing learning further. Pring (2013, p.191) supporting this in saying that "Learning builds on what has been learnt and provides a springboard for what is yet to be learnt." However, the number of topics in science does not always lend itself easily to this approach and ultimately could require reworking of the entire mapping of science teaching in secondary school.

The complexity of this task was clearly not a simple matter. Anderson and Sosniak (1994, p.200) found that forty years after Bloom's (1956) research the challenges had not been fully met and that a new framework should "make progress in solving the process-product relationship, facilitating the translation of goals into learning experiences and activities..." They referred to Bloom's desire to see his taxonomy applied to the learner (and others) in a productive way (Anderson and Sosniak,

1994, p.200). Whether this should involve breaking learning down to word knowledge is unclear. For example, Hoyle and Stone (2000, p.89-99) discussed “listening and speaking”, “reading” and “writing” with little mention of vocabulary development as a discrete activity. This could have been interpreted as a need to begin the development beyond individual words but since this was not clearly established the purpose of this thesis would allow for consideration of how the development of vocabulary might be appropriate in terms of learning in lessons. An approach that introduced new vocabulary ready for use with higher levels of learning following on during the lesson (or subsequent lessons) needed to be evaluated further.

Harrison and Morgan (2012, p.113) supported these concerns about language in saying “There is an increasing sensitivity to the challenges posed by the language of examination papers and of instruction in scientific subjects, especially for non-native speakers of English. It has been observed that in addition to technical subject-specific vocabulary, non-technical words such as instructional verbs have been sources of difficulty, and there are indications that other ‘ordinary’ English words cause problems.” This furthered the concern about the understanding of vocabulary even before considering the use of that vocabulary to interpret the requirement of its further use in the examination questions. However, the instructional verbs would not be considered here as they were subject to whole-school approaches as had become normal practise in developing examination technique. Their definition of “ordinary English words” is unclear and suggested the need to look further at language understanding beyond the science keywords, possibly in relation to the aforementioned socio-economic links. Given the focus in this thesis on the

knowledge of words and not their higher level use it would not impact upon the findings if other words also caused difficulty and further exploration into the breadth of vocabulary to be tackled could form a point of further study.

The question of what would constitute ordinary language was developed further by Lee and Fradd (1998, p.12) who also raised the concern that "...little attention has been given to the attainment of educational equality for all students..." They proposed (Lee and Fradd, 1998, p.18) that "...an instructional approach is needed that establishes congruence between the nature of science and the language and cultural experiences of the students." They added that (Lee and Fradd, 1998, p.19) more research is needed in bilingual and multilingual classrooms, suggesting the need to consider the variety of cultural contexts in studying science. This again highlighted the need to consider language variation between pupils and what constituted ordinary language could be highly variable according to the background of students.

However, the focus on the simplistic knowledge of words as opposed to higher level use of the words drew concern from Svensson et al. (2009, p.205) who claimed that "Empirical results show that frequently the meaning of expressions used by students in expressing their understanding of subject matter does not correspond to the meaning of those expressions in the subject matter theory that the students are expected to learn." He continued to say that "There is also often a lack of identity of meaning between the same students' use of the same expression from one place to another, even in very similar contexts." This highlighted the concern that developing knowledge of words might not correlate with their correct use in further learning.

Misconceptions and dual meanings of words would need to be considered as potential issues when devising the vocabulary learning tool and any assessment of success. Braund (2008, p.10-11) supported this, commenting that “The same terms of ideas are used to mean quite different things in different lessons.”

2.3.3 Previous experimental research on vocabulary

Mason, Mason and Quayle (1992, p.342) provided an example of where language had been claimed to have had a significant impact upon results at GCSE level. They developed a language course that they claimed impacted upon GCSE result increases from 30.3% to 55.6% (percentage of pupils gaining 5 or more A* to C grades at GCSE) in two years. Use of the York Language Aptitude Test comparing a control group with an experimental group showed a mean increase of more than double in the experimental group (Mason, Mason and Quayle, 1992, p.348). The experimental group had completed three workbooks. The first concerned grammar, word roots (including Latin and Greek roots) and language varieties. However, a valuable point was raised by Pampush and Petto (2011, p.9) who brought this link into question as they concluded that “there was only a weak positive correlation between the performance on a quiz of Latin and Greek medical terms and the students' performance on regular assignments.” Whilst this was based on 446 students (52 in more detail) and whilst only based on one course at University level it suggested that the issue of language was more complex than an ability to deduce meaning with pre-existing language skills where words were unfamiliar. Mason, Mason and Quayle said that their second workbook involved reading for learning. This was a common approach in science lessons so was not an obvious area for

further development. The third workbook involved writing for learning which was less developed in science and appeared to tie in to the wider ideas of balanced language development. Impressive claims were also made using reading and writing tests that were devised by the school (Mason, Mason and Quayle, 1992, p.350). This appeared to support the idea of a link between English language skills and examination results but did not provide any obvious insight as to which area of language skill development was most useful. Indeed it possibly indicated that a more coherent approach utilising various skills would be needed.

Lara-Alecio *et al.* (2012) produced one of the most directly relevant pieces of research relevant to this thesis. The background to their work in Texas, USA was strikingly similar in places to the context in the United Kingdom at present, demonstrating further that the issue of vocabulary in science was an international one (Lara-Alecio *et al.*, 2012, p.987-9). They identified ELL (English Language Learner) pupils as common in Texan middle schools (Lara-Alecio *et al.*, 2012, p.987-8) and accordingly carried out an experiment involving “A total of 166 treatment students and 80 comparison students from four randomized intermediate schools.” Whilst EEL/EAL pupils are not the focus of this study there would be EEL/EAL (English as an Additional Language) pupils involved. Also, Lara-Alecio *et al.* (2012, p.994) included low socio-economic EELs in their study. The aforementioned approach to word knowledge as a starting point in developing use of vocabulary was also in keeping with a second-language approach. Wallace (2012) further inferred a link between social class and language skills furthering this concept, though it was not directly explored in the literature.

Further analysis of Lara-Alecio *et al.*'s (2012, p.997) report identified the topics that were taught during their period of research. These included physics topics that appeared to be similar to Year 10 EdExcel physics GCSE specified topics that were used in this study. Lara-Alecio *et al.*'s (2012) positive results in these topics supported the idea that a successful approach could be developed utilising their research interventions. Those interventions included a variety of approaches including professional development for teachers, inquiry-based learning, integration of reading and writing, integration of technology, homework activities and mentoring. The interventions also included "direct and explicit vocabulary instruction" (Lara-Alecio *et al.*, 2012, p.996). Student-friendly definitions were included as one of those techniques that fall under this instruction methodology. It was not possible to deduce whether the vocabulary instruction was of value alone but this added value to the importance of understanding the role of vocabulary and whether it could be of use in directing future research towards other areas or a variety of approaches in combination. Lara-Alecio *et al.* (2012, p.1004-5) did however conclude that "the results of our study underscore the importance of implementing direct and explicit vocabulary instruction." They noted that they planned (Lara-Alecio *et al.*, 2012, p.1004) a larger study and that questions remained over the effect size of other influences on the variance in pupil results. Ultimately pupils in the treatment groups statistically outperformed their peers in 3 out of 5 tests. This supported the use of their research in developing the intervention used in this study.

August *et al.* (2009, p.345) carried out a preceding study in Texas using an intervention to develop "science knowledge and academic language" of middle school pupils. Their intervention consisted of professional development for teachers

and instructional materials to use with two units of study. The materials built upon a skill-based hierarchy not dissimilar to Bloom's (1956) taxonomy. Explicit vocabulary was included with 15 words per week selected from science and general vocabulary (August *et al.*, 2009, p. 354). Their research included both ELL and English speaking learners (August *et al.*, 2009, p.345). They stated an objective (August *et al.*, 2009, p.349) as "...successful interventions must be effective for ELLs but must not disadvantage English-proficient students because English-proficient students and ELLs are most often together in the same classrooms in the middle grades." They (August *et al.*, 2009, p.369) did note that the sample size was relatively small but used a selection of descriptive and statistical analyses to thoroughly examine the collected data. They concluded (August *et al.*, 2009, p.371) that there was a significant improvement when using measures of pupils' vocabulary in science, noting variance dependent on the pupils' teachers. This further suggested that elements of their approach could prove successful if reapplied.

This research formed the basis of the methodology and design in Chapter 3. It also indicated that interventions in science for vocabulary had been significantly effective for both EEL and non-EEL learners. Lara-Alecio *et al.* (2012) used this as one of the foundations for their own work. It was interesting to note that where August *et al.* (2009) opted to take account of teacher effect though finding relatively little variance, Lara-Alecio (2012) chose not to attempt to do this. This added weight to the decision not to evaluate teachers, particularly when combined with the ethical issues of doing so. The research required the good will of colleagues and subjecting them to a rating of their teaching would not have been conducive to goodwill and might have

resulted in placing them under unnecessary stress, even if they still had consented to participate which could not have been presumed.

These two robust trials of interventions showed a measurable positive effect on vocabulary and to a lesser extent science in general, and not just for pupils who are learning English as a second language. The case for a direct-vocabulary tool was not clear given the multiple variables in the interventions but such a tool was present in both pieces of research. This called for further understanding of whether the vocabulary element of the approach would be useful alone and indicated that if this was not found that a multi-faceted approach could be a stronger area of research in future. Either result would in a sense narrow the gap in the research by either indicating a need to develop the research into vocabulary or move back toward multi-layered (or alternative) approaches. Further research linking vocabulary to results would also be required to extend any conclusions from vocabulary development to examination results and general improvements in science.

Mohan and Slater (2006, p.303) supported this conclusion in saying “Few studies have taken a linguistic perspective to this theory–practice issue. In this paper, we explore how a science teacher socializes his students into the science register—thereby teaching students science language and meaning—which links what they do and observe with the theory being taught.” They discuss examples of language in science at length, concluding (Mohan and Slater, 2006, p.314) “Implications for the observation of teachers’ work call for a greater understanding of the development of the science register in science classes.” Mohan and Slater summarise the need to understand better the development of language in science through use of

vocabulary. This research met this demand, though allowed for further expansion of this understanding.

The literature appeared to support the conclusion that a direct-vocabulary based intervention in science as part of wider intervention could yield a statistically positive intervention effect in the region of 0.2 ($p > 0.01$) for science tests with slightly higher results for reading (based upon Lara-Alecio *et al.* (2012)). It further supported the conclusion that this effect may not be limited to pupils who are learning English as a second language. Furthermore there appeared to be a gap in understanding the role of a direct-vocabulary intervention as opposed to multi-layered approaches.

2.4 Developing language in the classroom

2.4.1 Approaches by other researchers

The literature thus far showed that a direct-vocabulary intervention had played a role in a wider suite of interventions to develop pupils' ability in science. In attempting to create an intervention that dealt with the vocabulary aspect alone it was therefore necessary to consider whether such an approach could be successful and what form it would take. The successes of the interventions by August *et al.* (2009) and Lara-Alecio *et al.* (2012) could have been rooted more strongly in the other elements of the intervention or in more complex interactions between those elements. However, the depth of the approach in these projects led to clear difficulties in developing an intervention and identifying what was responsible for the success so it was still deemed important to avoid a more wide-ranging intervention in favour of researching

the vocabulary aspect alone. However a deeper understanding of the methods used by this prior research was required in order to develop a vocabulary-based methodology that might build upon previous successes.

As previously, described, August *et al.* (2009, p.345, 352) developed an intervention which they named QuEST (Quality Science and English Teaching) and described as consisting of resources for teaching including "...instructional materials including a teacher guide and instructional charts, a student guide and instructional charts, and supplies for hands-on science activities..." It also included CPD [continuing professional development] activities for staff. This was clearly very thorough in approach but also a heavily managed approach. Time demands in school meant that extensive CPD and more particularly extensive management and monitoring needed to be avoided unless their use could be strongly justified. More successful approaches had been more explicit to teachers through supportive resources, guidance and brief (but sometimes frequent) CPD and support. This approach had helped to ensure consistency and coherence in the approach, as well as to sufficiently equip staff to participate. So the resource and guidance concept of QuEST and the important of training staff were retained but the more rigorous monitoring tools, for both teachers and for the researcher as line manager, were not utilised. This opened up the potential for more variation to occur between teachers and their classes and this was apparent in the later analysis. Ultimately the opportunity for more consistent approach was not taken since the more varied approach used would be more in keeping with the unmonitored use in future of such an intervention. The results were therefore more likely to reflect the realities of teaching. This approach would also allow teachers to continue to utilise their own

methods of engagement and motivation that the literature had found to be important in science. It would also allow for teachers to assess and take account of prior learning which had the potential to be highly variable between classes and pupils individually.

Lara-Alecio *et al.* (2012, p.995-6) also described their intervention as consisting of on-going teacher development alongside integrated resources for teaching. Again this was a similarly heavily-managed approach but did demonstrate potential in the technique. The benefits of providing resources seemed clear whilst the difficulties in doing this in heavily managed approach led to the need to utilise a less intensively managed and more teacher focussed approach. It was also unclear to what extent in this or August *et al.*'s (2009) trials the teaching resources had influenced teaching beyond language development and the descriptions seem to suggest a more wider-ranging influence.

The vocabulary element of August *et al.*'s (2009, p.354) work involved the aforementioned teaching of about fifteen words per week of science and general supporting vocabulary. Words were selected based on frequency in academic text and words deemed essential for understanding of key concepts. Glossaries were provided with simple definitions and visual images. This was supported by guided reading with activities to support the development of the vocabulary. Lara-Alecio *et al.* (2012) developed work by Carlos *et al.* (2004, p.193) which focussed on bilingual (or monolingual) pupils who were mostly of "working class Mexican American background." Ten to twelve target words were introduced each week alongside 30 to 45 minutes instruction per day and a review week every fifth week. Focuses

varied each day but included increasing depth of word use and various activities for word use. Lara-Alecio *et al.* (2012) used an approach of introducing keywords at the beginning of lessons then using them in different ways during the lessons. So the introduction of around a dozen words per week had been used in several studies and was accompanied by various word use activities and support. This provided an informed background for development of the method (in Chapter 3) but did include work outside of science and with pupils who were bilingual. The successes of Lara-Alecio *et al.* (2012) were built upon work with ELL/EALs but results suggested that use of work with English language learners could also be applied to a wider spectrum of learners.

Further discussion about the selection of the words to be used in this study came via Wellington (2000b, p.168) who suggested a taxonomy of words: Level 1 naming words, Level 2 process words, Level 3 concept words, Level 4 mathematical words and symbols to start to tackle this issue. He added caution (Wellington, 2000b, p.170-1) saying "...beware of meaning at higher levels..." asking whether pupils were ready given their language development and the need to teach for shared meaning. This led to these levels being used alongside a perception of the researcher as to which words might be best included. Level 3 words were included with those Level 1 and Level 2 words likely to cause issue from prior experience.

Wellington and Ireson (2008, p.175) also suggested use of a dictionary or glossary and highlighting new words, in keeping with August *et al.*'s (2012) use of a glossary. Dictionaries were available in the laboratories and glossaries existed in some textbooks though they would not have been available every lesson or widely utilised,

leaving a further avenue for analysis. Furthermore Wellington and Ireson suggested (2008, p.177) that pupils practise writing their own material, tying into the discussed concepts of using words as once they had been introduced. This was developed as an activity as part of the intervention.

2.4.2 The “Four Strands” approach to vocabulary development

The nature of the activities used to develop vocabulary was discussed at length by Nation (2007). The value of Nation’s work was summarised by the author of the “Confused Laowai” blog:

“To me the four strands of language learning is a succinct balanced approach of looking at language learning. Like I said, I’m not a good language learner, especially when I now look at the four strands and how I easily neglect some aspects.”

(De La Rouviere, 2012)

Nation, 2007, p.2) said that there was a need to be wary of using common sense to prioritise quantity over quality. He went on to say that “there is something about each of the language skills of listening, speaking, reading and writing that makes them different from the others.” He argued (Nation, 2007, p.2) that “It is thus necessary to give attention to each skill to make sure that these unique features are learned.” In support of this Waring and Takaki (2003, p.131) say “The general picture that emerges from these studies is that learners do learn vocabulary from their reading but not very much.” They quantified this saying “In many of the studies,

typically the gains in scores after reading are only just significant and not much better than random guessing on the tests.” So the intervention should not only familiarise pupils with reading keywords but go further so as to enable the pupils’ deeper understanding of these words. Prioritising quality of that learning would be more important than repetition. This led to support for an intervention which spent adequate time explaining keywords rather than one which frequently revisited them for less time. However, Nation (2007, p.2) pointed out that “It is also possible to distinguish accuracy from fluency and thus see the necessity for giving fluency practice for each of the skills.” This suggested that some level of practise and repetition was still important. Revision and the building of topics (to a lesser extent) would help to achieve this but the full impact of revision was difficult to measure as it might be deemed to be best measured at the point of the actual GCSE examinations which would not be appropriate due to the other pressures of those days on the pupils.

Nation’s (2007) “Four Strands” explained an input/output approach to language development. The four strands identified were (Nation, 2007, p.2) “meaning-focused input, meaning-focused output, language-focused learning and fluency development.” Nation went on to say that the tag of “strands” is used to emphasise that these forms of language development run throughout a language course and are not limited to certain points within the learning. This meant that the use of language would need to expand beyond the intervention and become embedded in the whole lesson. Lara-Alecio *et al.* (2012, p.996) also presented the importance of an output driven approach to their intervention.

The conclusions by Nation (2007, p.10-11) are summarised below:

1. Reading and listening activities should be used;
2. Highlight and explain words as they arise;
3. Encourage writing and speaking in various genres and situations;
4. Use group work tasks for reading and writing of various forms;
5. Use teacher-led intensive reading helping learning to tackle patterns, items, grammar and discourse;
6. Train learners in language strategies;
7. Provide fluency activities in listening, speaking, reading and writing;
8. Provide an equal balance of listening, speaking, reading and writing;
9. Use repetition;
10. Tailor to learners' needs using analysis, monitoring and assessments.

The intervention was developed keeping these conclusions in mind and each of the points was met to a varying level either through the intervention or the normal approach of teaching alongside it. However the ability to use monitoring and assessments was not provided to its fullest extent in the study due to time limitations though the information was utilised beyond the study in student examination preparation.

Funk (2012, p. 310) supported the variety in Nation's (2007) conclusions, in saying that "Language educationists and educators need to consider a number of state-of-the-art approaches and models..." He also stated that consideration should be given to "...possible impact on curriculum planning, distribution of learning activities, and progressional planning for textbooks and classroom instruction." Distribution of the

activities linked again to the idea of developing language through higher level skills. Curriculum planning and progression along with resources such as textbooks were considered key to this study due to the limited available time and forthcoming GCSE examinations. These issues had to be carefully balanced with time dedicated to an additional activity. However, it was also important to evaluate existing and potential activities rather than dismiss them because of these pressures. This added weight to the need to evaluate individual elements from the outlined approaches to ensure time and focus is spent on those parts that are most effective or to reinforce the need to use a variety of approaches which might not individually show a clear link to pupil development.

Nation (2007, p.2) also advocated that "...the learners' main focus and interest should be on understanding, and gaining knowledge or enjoyment or both from what they listen to and read." He (Nation, 2007, p.2) suggested that typical activities include extensive reading, shared reading, listening to stories, watching TV or films, and being a listener in a conversation. These activities would be longer than would allow in a simple intervention but highlight the importance of an awareness of language use across the range of classroom activities. Nation (2007, p.2) did emphasise other more general learning conditions that were essential for learning to take place through input. These included prior familiarity with most of the listening or reading, interest from the learners, who must want to understand the work, less than five words per hundred unfamiliar to learners and background knowledge to aid learners. He said that similar conditions exist for output but included the importance of strategies (such as availability of dictionaries) to support output and outlines the importance of speaking and writing. The combined message here was one of the

embedding of literacy into the lesson as a whole. However, these points did not clarify whether an intervention could be used where current teaching practise appeared to leave pupils with vocabulary gaps.

Wellington and Osborne (2001) supported the use of spoken language in science teaching. They did however warn that (2001, p.39-40) “teacher-pupil exchanges alone...are unlikely to develop...oral skill” and recommended structured and well-disciplined pupil-pupil talk. Such activities would contribute both input and output for pupils in balanced measure – meeting Nation’s (2007) recommendations. Mason, Mason and Quayle (1992, p.347) supported the importance of pupils working together independently and in pairs.

Wellington and Osborne similarly supported reading as an activity (2001, p.59-60), highlighting that coaching and practice were needed to learn the skill of reading in a critical and active way. Wellington and Osborne (2001) provide lengthy further discussion on writing techniques too. Overall much of the detail of Wellington and Osborne’s (2001) writing supports the summary and research of Nation (2007). These findings added extra weight to a varied approach and one which utilised the “four strands” in a balanced way. Further support came from Mason, Mason and Quayle (1992) as their aforementioned study is based upon workbooks which focus on language, reading and writing.

2.4.3 Vocabulary development activities

More precise detail was considered for the activities which would meet the “four strands” requirement. Knipper and Duggan (2006, p.462, 465-468) found that writing to learn was not the same as learning to write and suggested a variety of techniques and activities for using writing as a tool for learning. These were often short activities, carefully structured and framed, emphasising the importance of using language in a structured and managed way in the classroom. These findings again supported variety but with careful consideration of the structure and content of the activities.

Traianou (2012, p.217) did raise concern in pointing out that the whole notion of a necessary acquisition of a set of skills and knowledge is cultural yet at the same time it is accepted that construction of personal experience is a key part of making knowledge useful and understood. This links to the findings of Lee and Fradd (1998) previously mentioned. This point suggested a link between cultural background and the whole nature of language that was complex. This also highlighted a danger in considerable over-simplified approaches where the pupils would all be different and might respond in different ways. This could have been seen to further support variety in approach so as to cater for the variety in pupil cultures and responses by the class teacher.

OFSTED (2013, p.15) added some further points of consideration for use when designing the intervention. They identified those areas which they believed generate misconceptions that prevent “good” teaching. These areas of consideration were

(OFSTED, 2013, p.15): pace (pupils should be concentrating as opposed to rushing); the number of activities (quality in fewer rather than too many); over-detailed and bureaucratic lesson plans (detracting from the focus on pupils' learning); inflexible approach to planning (no one structure will always be suitable); and constant review of learning in lessons (allowing time for pupils to complete work before reviewing). The intervention would ultimately take account of these points though time constraints again would have an impact on how well these objectives would be met. Teacher flexibility was at the forefront of these considerations.

In designing a vocabulary intervention there were to be different types of language (or word groups) to be considered. The literature supported the careful use of reading, writing, speaking and listening and in roughly equal proportion. Variety in the activities chosen would be important. Working in pairs and independent work were also found to be important in enabling language development, not least as they would enable balanced output and input activities to be designed for the pupils. The nature of development was highlighted as a complex issue potentially requiring multiple approaches to ensure success. Within any approach the ability of a teacher to be flexible to the needs of a particular class or students was very important.

2.5 Testing knowledge of vocabulary

A standardised test was not available for such a specific set of vocabulary so the approach used by Mason, Mason and Quayle (1992) would not be possible until the potential point of widening analysis beyond specific vocabulary towards wider literacy measurements. As such a bespoke test was developed. Waring and Takaki

(2003) developed a test for learners of a foreign language in order to learn which words they had learnt. They chose 25 words to make up their study and discussed their need for a 3-stage test in order to understand different levels of word knowledge. They ultimately decided to test for word recognition, using a multiple-choice meaning section (with an “I don’t know” option) and a meaning test. Their results were given for each of the three sections and whilst scores declined in each section they were indicative of each other, meaning that any of the three tests could be useful for simple comparative purposes. Prior experience suggested that the approach Waring and Takaki (2003) used in Section A (i.e. ticking familiar words) would prove unreliable. Pupils might have ticked any words or been mistaken in their understanding or familiarity with words, particularly if their reading skills are weak and they misread words. Section C (defining words) was likely to have been both slow and difficult to answer as well as to mark afterwards. The pointed towards a use of a test in keeping with Section B – a multiple choice test. However, the multiple-choice test required careful development due to the potential issues in constructing the various options for each set of choices (Cantor, 1987).

Cantor (1987) recommended against using a “Don’t Know” style answer as a “distractor” in a multiple-choice test but knowing where pupils genuinely did not know an answer was deemed important enough to override Cantor’s concerns, especially given that Waring and Takagi (2003) had successfully used this approach. Cantor (1987, p.88) provided a checklist for reviewing questions which was utilised. Campbell (2011) also provided a brief description of good practice, highlighting the need to consider the test in light of its use – for example querying those questions

where pupils who score highly overall appear to have difficulty with. The test was evaluated as part of the quantitative analysis given this recommendation.

In developing measurements for testing the common quotation concerning testing used by President Barak Obama concerning education became worthy of consideration:

There's a saying in Illinois I learned when I was down in a lot of rural communities. They said, "Just weighing a pig doesn't fatten it." You can weigh it all the time, but it's not making the hog fatter. So the point being, if we're all we're doing is testing and then teaching to the test, that doesn't assure that we're actually improving educational outcomes.

(President Obama, 2009, online)

The point being raised here was that in developing testing, retaining clear perspective of the ultimate objective of improving science education was vital to avoid becoming focussed on improving test scores as opposed to science learning. The test could prove an indicator of the latter but this required ongoing evaluation.

2.6 Conclusions from the review of literature

The reviewed literature pointed towards a need to be concerned with the complexity of language in science. It also suggested complex issues surrounding the effects of teaching interventions upon pupil attainment. In particular the literature identified a need to consider pupil motivation and engagement as key parts of the puzzle that

might unlock further achievement. There was clear support for the use of a variety of activities and also variety in the types of literacy activity used. The research was sufficient to suggest ways of tackling literacy and specific vocabulary difficulties. Furthermore two studies had provided templates for a research design and method, though they needed to be reviewed for application on a smaller scale and in an English school setting. Nevertheless they gave support to the idea that tackling issues with vocabulary within a bigger picture of effective teaching and learning could be successful in helping pupils understand language and so engage in their science lessons.

There were a few key evaluative points that became apparent during the review. First of all, Gorard, See and Davies (2011, p.120) said that much of their work was based on research “conducted in the USA.” Nation’s (2009) work was conducted from New Zealand, whilst Lara-Alecio *et al.* (2012) and August *et al.* (2009) were researching in the United States. This highlighted two contrasting points: the most prominent perhaps that the issues being raised in the literature were being experienced in other countries teaching science through the medium of English. The contrasting point would be that there were differences within the nature of the language and context of the teaching between countries. This meant that this study would be useful beyond English schools but that there would be inherent differences to take account of before presuming that any successful interventions in England would be successful elsewhere. Given the variance between schools even in England this would need to be regarded as the case even within a more confined geography.

Secondly, Harris (2011, p.385) argued “that while teachers focusing on behavioural and psychological engagement could easily describe the successes of their lessons in affective terms, identifying concrete learning objectives was more difficult.” He (2011, p.385) went on to say that “This indicates that overemphasis of affect and participation may cause teachers to unwittingly shift their focus away from the cognitive engagement and academic learning that education should be trying to promote.” This made a second point of interest as it flagged up the dangers of developing teachers to use a system rather than develop engagement as a whole, something which the literature would warn against. Wallace (2012, p.298) concurred with this comment, saying “Studies have also been conducted on the detrimental effects of highly specified standards on children’s desire to learn science.” Harris (2011) and Wallace (2012) furthered this issue in their work on educational research and development of practice. They emphasised that change must be sustainable and focussed on developing learning not on creating tick-the-box exercises which were not focussed on the correct outcome.

A third point was that as part of the research process there would need to be effective professional development of staff. Reay (2006, p.303) said “...initial teacher trainees are left ill-informed and, I would argue, ill-equipped to broach, let alone tackle, the greatest problem the education system faces: that of working class educational underachievement, alienation and disaffection.” This suggested a wider impact of the research in that it could help to tackle an issue in teaching and involve staff in the process. This would be a good model for tackling the problem identified by Reay and the research could also further the call for better training on literacy and language in science, particularly as the review makes clear the possible links

between socio-economic status, reading ability and science. Those links are not always direct and clear qualitative experimental measurement of them was limited but identifiable.

As a final point, Lakin and Wellington (1994, p.175-190) described their research into teachers' perceptions of science teaching and ultimately concluded that innovation in science teaching may not come from teachers and may even be impeded by them. Research would ultimately involve other teachers and success in this approach could be a model for overcoming difficulties in involving staff in developments in science teaching.

These conclusions helped to shape the research design and method so that it tackled the gap in the literature but also developed research that had shown success in practise not just through theory. It also took account of the literature on literacy and vocabulary in designing an intervention. The research design and method also paid heed to the warnings about narrow and over-simplistic attempts to find impacts and the need to involve staff in CPD and the project as a whole. This included avoiding "tick-box" approaches and aimed to improve teaching and learning in a wider sense, supported by an intervention and not dominated by it.

Chapter 3 – Research Design and Methodology

3.1 Introduction

A general discussion on the selection of the two design frames that were utilised, (quasi-experimental and survey) is followed by sections with more detailed description of the individual design frames. Ethical considerations were embedded within these sections and methodology was outlined after each research design frame.

3.2 Selection of research design frames

Verma and Mallick (1999, p.10) noted that “research has been classified in various ways...” Within their classification this thesis was considered to be applied or field research (Verma and Mallick, 1999, p.11). This was useful for an awareness of critical differences between types of research; with this research discrete from pure research, action research and evaluative research. This was important in that control of all factors was difficult due to the emphasis on the effects on particular groups of students. This meant that there were various complex variables that required careful consideration of the appropriate research frames in order to ensure that these variables were controlled or accounted for as far as possible in order to draw valid conclusions. There was also an early decision not to conduct the research via action research and that the aim was to evaluate the effect of a vocabulary intervention and it did not aim to develop this further at this stage.

Selection of the frames used was based upon several factors. A true gauge of pupils understanding of keywords before and after the vocabulary intervention was desirable in order to investigate the level of development that had taken place. This could not be carried out by an experimental design due to ethical considerations surrounding the use of control groups so a quasi-experimental method was used instead. However, the complexity of the findings would require more insight than numerical analysis would provide and so a design that would delve into the experiences of pupils and staff was required. As this insight was concerned with experiences of the intervention a survey post-intervention was deemed an appropriate approach. Also, the research built upon the studies made by August *et al.* (2009) and Lara-Alecio *et al.* (2012) and so consideration was given to these research designs and methodologies. However, both of these studies were carried out using a much broader range of schools, teachers and pupils. This meant that the design needed to be adapted for use over a shorter period of time and with a smaller sample size, which further supported a mixed methods approach with a quasi-experimental element.

The concept of this approach agreed with Thomas (2009, p. 74) who said that research paradigms need not be seen as conflicting or independent but actually should be used as appropriate to the question being considered. It also acted to counter the concerns of Hargreaves (1998, p.48) who pointed out that the social science researcher must ascertain “that his [or her] perception and analysis of a social system does bear a close resemblance to what is actually occurring.” Hargreaves (1998, p.48) suggested that “...observation and interview do not in themselves adequately safeguard our need to support what is claimed...” Freebody

(2003, p.133) concurred with this, saying "...what is said about a certain phenomenon or event in interviews cannot be taken as a reliable proxy for the observation of that phenomenon or event." In other words there was a need to validate perceptions in interviews using multi-method approaches if the outcome was to be indicative of what was actually happening as well as what people perceived to be happening. Neither statistics nor survey alone could give a complete and reliable picture.

In following this approach Walford's (2001, p.85) claims were heeded in that "...many research reports increasingly use the data generated in interviews as the sole, or major, source for such descriptions." He remained unconvinced that this was appropriate, saying (Walford, 2001, p.85) "...most research is more interested in what people do than what they say they do" and "...what people say when they are interviewed should be treated with extreme care". The intent here was to be able to integrate the factual observations of progress in vocabulary with the perceptions of the participants. With careful analysis it was possible to differentiate between the perceived learning reflected in the statistical analysis and the benefits felt by the pupils and staff, as well as perceived problems and potential further improvements.

Thus the research questions are presented in Table 3.1 below with the selected principal design frame.

Table 3.1 – Research questions and principal design frames

<u>Question</u>	<u>Principal Research Design Frame</u>
1. Do pupils show an improved ability to define keywords following the introduction of a pedagogical intervention in during a GCSE physics module?	Quasi-experimental
2. What worthwhile benefits do pupils perceive from the change in pedagogy?	Survey
3. What worthwhile benefits do teachers of the pupils perceive from the change to pedagogy?	Survey

The first question most closely mirrored Lara-Alecio *et al.* (2012) and August *et al.* (2009) in seeking to measure the effect of an intervention. In order to do this and take account of Gorard, See and Davies (2011) in seeking to measure the effect size it was necessary to conduct this research in a positivist (or scientific) form though supported by the findings of the survey to overcome the difficulties in drawing conclusions from the small sample with complex variables involved. Questions two and three then enabled any measured effect to be examined in relation to pupil and teacher experiences using the findings of the survey in combination to a lesser extent with reflection from the quasi-experiment where it became relevant. This

approach matched the recommendations of Hargreaves (1998), Freebody (2003) and Walford (2001).

3.3 Development of research design frames

The research was dependent on two different design frames that interlinked and yet were also carefully utilised independently to illicit robust conclusions. Thus the two design frames of quasi-experimental design and survey design are discussed in turn. The outline of the design frames followed a common approach with general definition and discussion of the nature of the design frame, followed by the identification of the population, the sample that was used from the population and finally the process that was followed in utilising the design frame, including sequencing.

3.4 Quasi-experimental design

3.4.1 Description of the quasi-experimental design frame

Cohen, Manion and Morrison (2007, p.272) explained the purpose of experimental investigation by saying that “If rival causes or explanation can be eliminated from a study then clear causality can be established; the model can explain outcomes.”

They went on to describe the process whereby “...investigators deliberately control and manipulate the conditions which determine the events in which they are interested, introduce an intervention and measure the difference that it makes.”

Thomas (2009, p.124) concurred, highlighting the cause and effect demonstration through use of controls. However introducing a control was problematic,

predominantly due to ethical concerns, and so a quasi-experimental frame was found more appropriate. Nevertheless the identification of variables and an awareness of those which should be taken account of, and possibly even controlled, was still a logical first step.

Table 3.2 below shows the experimental variables for question 1.

Table 3.2 – The experimental variables

Independent variable	The application of a vocabulary intervention (by means of a series of worksheets)
Dependent variable	Pupil ability to define keywords (by a multiple choice test of 50 of the words).
Control variables	The course being studied, age/year group, point in study of the course, time spent studying the course in school.
Extraneous variables	Teacher influences, peer influences, additional activities taken by pupils independently.
Intervening variables	Unknown but could have included motivation and aspiration.

This table continued to highlight the complexity of the variables involved in the study and is not exhaustive. The intervening variables were difficult to establish at this stage but, along with the extraneous variables, would to an extent be established

through the survey. This would allow more sound conclusions and thorough evaluation later. The control variables were accounted for in the research design and methodology but it was also necessary to consider whether the controls had been successful.

The hypothesis and null hypothesis were established using these ideas and useful for drawing conclusions from this study. They were:

H_1 = There is a positive correlation between the applied pedagogy and pupil ability to define keywords.

H_0 = There is no link between the applied pedagogy and pupil ability to define keywords.

Lara-Alecio et al. (2012) and August et al. (2009) used experimental studies in their research. Cohen, Manion and Morrison (2007, p.274) addressed the difference between experimental and quasi-experimental research in saying that “The single most important difference between the quasi-experiment and the true experiment is that in the former, case, the researcher undertakes his study with groups that are intact, that is to say, the groups have been constituted by means other than random selection.” Mertens (1998, p.77) used a similar description, saying that quasi-experiments were “‘almost’ true experimental designs except that participants are not randomly assigned to groups.” As is the case in the aforementioned reports it seemed that this form of experimentation is the appropriate one to this study, both for logistical and ethical reasons. Pupils were already placed in sets and it would not be practical or ethical to disturb their learning, which could impact upon their results

and relationship with their teachers, for the purposes of this study. This was successful but more difficult to establish variance due to the sets rather than pupil differences. The data was also used to establish variance over time rather than with comparison to a control group.

3.4.2 Ethical considerations

Ethical considerations were particularly pertinent given the involvement of pupils in the research and potential impact upon them. The acknowledgement of ethical considerations has grown in recent years in social research (Cohen and Manion, 1994, p.347). The main principle behind this growth was to “not harm the participants in the research” (Gorard, 2013, p.187). Gorard (2013, p.188-190) also pointed out that conflicts of interest could arise at a higher level and thus the impact on the school (and even Local Authority and other overseeing bodies) also had to be taken into account. Given that the literature suggested the use of literacy activities and development of vocabulary in science it seemed unethical to withdraw this from a control group even though this would make it more difficult to establish a comparative effect. Further ethical decisions were made as part of the development of the methodology.

3.4.3 Population and sampling

Gorard (2013, p.85) warned that effect size was often small in social science and there was a corresponding issue with the relatively small sample size in this study. Lara-Alecio *et al.* (2012) and August *et al.* (2009) measured a variety of effect sizes

indicating possible effects in a larger population in the region of 0.1 to 0.3, however attempting to replicate the findings would require a much larger population. Given the available access and time for this study this was not feasible. It would also suffer from uncontrolled variables due to variation in teaching styles not only in science but in whole school approaches to language and mathematics. Gorard (2013, p.86) suggested the use of Lehr's formula as a heuristic approach. This indicated that an effect size of 0.157 would need a sample of 649; well beyond the means of this study. However, larger effect sizes would be apparent with smaller samples and descriptive statistics would still prove useful in answering the research questions, ultimately validating a statistical approach.

Corte, Verschaffel and Van de Ven (2001, p.44) had proven the success of using smaller samples with a design studying 79 pupils in their experimental group compared to 149 in their control group. Their analysis of text comprehension strategies found several interesting results though they failed to be statistically significant (Corte, Verschaffel and Van de Ven, 2001, p. 52-57). The supporting interviews (Corte, Verschaffel and Van de Ven, 2001, p.55) demonstrated the value of a mixed-method approach given that the statistics fell short of significance but there was still meaningful development of knowledge within the study. The lead to an attempt to conduct a 100% sample of the population but analysis using Creative Research System's (2012) calculator suggested that there was potential for the sample to drop to 81% of the population, should absenteeism or parental/pupil withdrawal become a factor, with statistical effect sizes still a possibility (for 95% confidence level and confidence interval of 5). Ultimately 82 pupils (90.1%) formed the sample with the other nine pupils being absent at key points in the research.

Attempts to maximise the population size led to the consideration of including Year 9 pupils in the study. Pupils in Year 9 followed a similar course (differentiated appropriately by their teachers to take account of their lower starting point) but the different point in their development (age wise as well as through variation in teaching and learning each year) was deemed as being likely to have a significant impact upon the results and as such they would not be a valid part of the same population without attempting to account for these issues, which would have been difficult (if not impossible) to achieve satisfactorily. This would particularly have been the case if Piagetian development had been taken to be relevant. As such Year 9 pupils were not included in the research and it was limited to the 91 Year 10 pupils.

3.4.4 Group size

The population was retained, as per the quasi-experimental approach, in original teaching sets, as summarised in Table 3.3 below.

Table 3.3 – Pupil teaching groups

<u>Year 10 Set</u>	<u>Included?</u>	<u>Number of pupils</u>
Triple	Yes (studying the physics 1 module at a similar time)	26
1	Yes	24
2	Yes	20
3b	Yes	11
3b	Yes	10

The table shows a Triple Science group who selected to study the three discrete GCSEs of biology, chemistry and physics and had an additional five hours a fortnight of timetabled time in order to achieve this. Four groups studied the compulsory Core Science GCSE. Set 3 had been sub-divided into two smaller groups due to flexibility in timetabling that had occurred earlier in the academic year.

3.5 Survey design frame

3.5.1 Description of the survey design frame

As discussed earlier in this section, this study recognised the value of a mixed-methods approach. Corte, Verschaffel and Van de Ven (2001, p.55) used an evaluative approach to support their conclusions with success. They noted the key weakness of their quasi-experimental approach was that it could not distinguish the relative importance of aspects of an intervention. However they argued that it was

still an appropriate approach for evaluative purposes. This study replicated that approach though the time scale and scope of this study meant that the evaluation was conducted through a survey. Hargreaves (1998) successfully evaluated sets in a secondary school in this way too. Gorard (2013, p.129) stressed one of the key benefits of this approach with the earlier mentioned example for population size: “The in-depth work allows an understanding of the surprising results.” The mixed-methods approach armed the researcher with information that was useful when the unexpected occurred. Given the rich and varied comments that were not anticipated in the findings, this was shown to be an appropriate approach to this research.

Thompson (1999, p.1) spoke broadly of the use of surveys by educators for a variety of reasons across various topics. Thompson (1999, p.2) said that the identification of perceptions was a possible project that could be addressed within this design frame, though Thomas (2009) questioned whether survey was more a methodological approach than a research frame in its own right. Irrespective of whether it was truly a design frame or a methodology this supported the suitability of survey for answering the two questions posed in this research about pupil and staff perceptions of the intervention. It was decided to focus on ascertaining perception to support numerical findings of impact measured by the quasi-experimental approach noting that perception was not the same as observed reality. As such there was potential to have extended the study to include a further design frame for observing the research events as they unfolded, though given the researcher’s role within the department this did not seem necessary. Different conclusions might have been reached had the researcher had more distance from the department during the study.

Cohen and Manion (1994, p.83) described a survey as gathering data at "...a particular point in time with the intention of describing the nature of the existing conditions, or identifying standards against which existing conditions can be compared, or determining the relationships that exists between specific events."

This was in keeping with the intent of this element of the research. Sapsford (1999, p.1) however said more simply that a survey "describes a population". Even if this were the case then the descriptive would be used analytically. Sapsford's (1999, p.1) definition could have been interpreted as an overly simplistic approach, which this survey was not when considered in terms of the full analysis. The definition was nonetheless useful for establishing the purpose of this research frame; particularly the importance of knowing the population and sample. In particular there was a need to ensure that the sample was representative of the population.

3.5.2 Population and sampling

Population size was common to that for the quasi-experimental design and as such incorporated all 91 of the Year 10 pupils studying the GCSE Core Science and GCSE physics courses for the question that required pupil perceptions to be addressed. Seven teachers (including the researcher) were involved in teaching these pupils and so formed the population of six (excluding the researcher) for the third question.

Sampling was more complex than for the quasi-experiment given the need to ensure that a variety of views were collected. This needed to be balanced with the intensity

of sampling which meant that a large sample would not be practicable, nor necessarily useful if the data gathered had become too complex to analyse effectively. McAteer (2013, p.75-6) said of participant selection: "...there will be a range of viewpoints and perspectives on any issue and, as a researcher, you must try to ensure that voices from all these perspectives, if possible, make their way into your data set." It seemed appropriate therefore to ensure that the sample covered the range of classes (to ensure application of the tool by different members of staff was accounted for) and the impact of the intervention, based upon the results of the quasi-experimental test. The criteria for this quota sampling (Cohen and Manion, 1994, p.89) resulted in three groups for each set of pupils: those who experienced the strongest positive (or least negative effect), a middle-effect group and a least (or negative) effect group, based on post-test scores.

3.5.3 Group size and composition

Previous research had shown success with various group sizes for survey samples, despite varying opinion on this (Krueger and Casey, 2000; Lewis, 1992; Watts and Ebbutt, 1988). It seemed appropriate to aim to have groups of around six as an approximate middle ground from the suggestions. However, flexibility allowing for practical issues common to working with pupils (such as return of consent forms, availability and attendance) meant that conducting groups as pupils were available was much more important than ensuring a particular group size. This worked successfully. Similarly practical matters meant that the particular pupils in the individual groups were not subject to control, only the complete sample.

3.5.4 Ethical considerations for the survey

The survey research frame involved more complex ethical considerations than the quasi-experiment due to the in-depth interviews and the involvement of children and colleagues, which resulted in data which had potential wider potential impact upon the participants and the host school than the test scores. Whilst there were specific considerations for the pupils and teachers involved in the research, many of the points concerning participants are common to both.

The British Education Research Association (BERA) (2011, p.5) stated that the dual role of researcher-interviewer had the potential to “introduce explicit tensions in areas such as confidentiality”, something which Gorard (2013, p.188-190) also raised concern about. McNiff and Whitehead (2006, p.86) supported this concern, emphasising the importance of “protecting your participants [and] assuring good faith...” McNiff and Whitehead (2006, p.86) advised “Make sure you do not name or otherwise identify your participants, unless they wish.” Zuber-Skerritt further emphasised (1996, p.16-17) that permissions be obtained and that there is responsibility for confidentiality. Thus confidentiality was maintained for the interviews. Had a child protection issue or conflict with school policy arisen then these would have taken precedence but this did not arise. Data was stored securely to prevent inadvertent breach of confidentiality. None of these considerations had any detrimental impact on the study.

McNiff and Whitehead (2006, p.87) said “Always do what you say you are going to do.” Zuber Skerritt (1996, p.16-17) also talked of the importance of “open-process”

during research. In concurrence with this all participants were given information sheets about the research with their consent requests. That information (see Appendix C) included a clear, honest and encompassing overview of the research aims. This did mean that participants were aware of the focus of the research and there was potential for this to sway comments in favour of (or at least towards) keywords where it might have been useful to measure whether participants did this without knowing the main focus. However, ethical considerations were deemed more important, particularly given the level of good will and trust required not only for the research but for the ongoing work of the science department in question – reinforcing BERA (2011) and Gorard's (2013) aforementioned concerns. This remained a potential weakness of the research design frame but not one which impacted upon the main focus of the conclusions and responses to the questions.

Zuber-Skerritt (1996, p.16-17) added that those who did not wish to participate should have their views respected. Whilst no pupil or teacher directly declined to participate, the letter to all participants (see Appendix C) made clear that they were not compelled to participate nor would there be any form of reprisal if they did not want to be involved. A period of 14 days after the interviews was allowed for withdrawing from the research. No participant made the decision to withdraw so this again did not impact on the research data and meant that participants had effectively further endorsed the use of their comments by not withdrawing them.

A final general consideration for the focus group interviews was raised by Cohen and Manion (1994, p.369-370) who raised the risk of deception in interviews and Sikes (2000) who documented examples researchers finding that participants had not told

the truth during interviews. There was no solution to this possibility other than to remain aware of the possibility when drawing conclusions. The nature of this study meant that it would have only been likely to generate deception if participants felt there would be reprisals or concern over responses. Given the collaborative, open and friendly manner of this research, deception was not obviously encountered but remained a possible error in the final conclusions. The use of quantitative test data alongside the qualitative helped to provide rigour to any conclusions made about the success of the intervention.

3.5.4.1 Ethical considerations for children as research participants

Consent for research with children needed to come from the legal parent or carer for each child (McAteer, 2013, p.57; Borg and Gall, 1989, p.94-100; Wellington, 2000a, p.125). Thomas (2009, p.150-1) supported the opt-in approach, particularly for children, pointing out that “There is an unequal power relationship between any adult and any child.” As such pupils were provided with copies of the letter and information found in Appendix C – affording them full involvement in the consent and understanding of the aims of the focus groups and research but requiring parental consideration/carers and consent. This led to issues when pupils did not gain consent, in one instance due to language barriers and in three cases failure to return the form for undisclosed reasons. The research therefore suffered from not being representative of pupils with a different language background at home and of pupils who did not return consent forms. It was not clear whether the importance of ethical approach (and constraints that resulted) skewed the data significantly. Unfortunately the complexity of resolving this was not feasible with the time and financial resources

(for translation) available. However, the pupils appeared to reflect on general experiences so the impact of this issue might not have been significant (though this cannot be absolutely presumed).

Cohen and Manion (1994, p.373) cited several examples of ethical dilemmas including one on a study into girls' under involvement in science and technology – highlighting how the issues can arise in the surrounding research area of girls and science. Cohen and Manion (1994, p.373-4) suggested that survey methods and the “evaluation of developments” are particular problem areas for inexperienced researchers. Walford (2001, p.137-8) exemplified the issues that can arise if an interview leads to an uncomfortable point of view being raised. These points led to preparation must be made for potentially uncomfortable comments that might impact on the interview – in terms of social, moral and cultural duties particularly. Safeguarding was held as being of absolute importance but no issues of concern were raised so these worries did not impact upon this research.

BERA (2011, p.7) specified that “offering sweets to school-children” would be an inappropriate reward. The difficulties in giving an appropriate reward and its perception by both those partaking and those not asked to partake meant that no reward was offered for what were relatively short focus groups in keeping with those regularly held by various staff members in school.

3.5.4.2 Ethical considerations for teachers as research participants

Whilst this research did not evaluate the teachers involved, the nature of the study did mean that variation between groups/sets and thus the teachers involved become apparent to an extent. Cohen and Manion (1994, p.376-378) discussed ethics surrounding teacher evaluation but focussed predominantly on more formalised assessment of teachers. Anonymity remained important to avoid unintentional evaluation being made.

Teachers were provided with refreshments in meetings (including cakes) but not as a direct incentive or thanks for attendance, with this being a normal habit between staff at various meetings in the school. Teachers were not coerced or pressured and very willingly gave of their time.

3.5.4.3 Ethical considerations of access and conflict of interests

McNiff and Whitehead (2006, p.86) said that some of the main considerations concerned “negotiating and securing access...” This concern was important given the dual role of the researcher. Cohen and Manion (1994, p.354) said that “Investigators cannot expect access to a...school...as a matter of right. They have to demonstrate that they are worthy...” They advised contacting the relevant person (or people) of authority, followed by outlining the research to be conducted, including the ethical considerations (Cohen and Manion, 1994, p.354-5). As such the consent of the Head teacher was gained in writing via a written request, after provision of the required information and also an informal discussion about the research.

In response to Gorard's (2013, p.188-190) concerns regarding conflict of interest; neither this research nor the surrounding research activity within the University of Birmingham was funded in any way by the school. As such there is no question of finances placing pressure on the outcomes of the study.

3.6 Order of research

For the quasi-experiment a design suggested by Cohen *et al.* (2007, p. 283) was used, as shown here using the notation system that is now in common use and first proposed by Campbell and Stanley (1963):

Figure 3.1 – Order of research

O₁ X O₂

This sequence and approach was a pre-test, intervention, post-test approach (Cohen *et al.*, 2007, p.282-3) and was a simplistic but suitable approach given the small time frame of this study. If this study had been extended then multiple testing points might have been more appropriate in order to better evaluate the impact of varying aspects of teaching which might have been influenced changes in score, such as revision. O₁ was conducted in January prior to commencement of study of the physics 1 module. The intervention (X) took place across the period of study, prior to the repeat of the test at O₂ before the Easter holidays (at the end of March). This timing allowed for completion of the teaching of the physics module and effectively

almost a full term of use of the intervention to gauge effect. There was potential to estimate effect over longer periods of time as had been done by August *et al.* (2009) but in this smaller scale study this was not attempted without more extensive data. The time allowed was sufficient to gauge effect on a small scale and indicate whether further research would be useful.

Despite the weaknesses of quasi-experimentation, there were some huge benefits to utilising this process as discussed by Cohen *et al.* (2007, p.274). These included that observer effects had a bigger influence in other research frames and this approach would ensure that observer effect was minimised. The survey element of this thesis was designed to complement this and the elements that were difficult to control and those that become subject to the understanding of the nature of the individual student responses were better analysed and potentially understood.

As the survey evaluated impact it was deemed most appropriate to carry this out upon the completion of the use of the intervention but as soon as possible afterwards in order to ensure participants' memory was as recent as possible. There was no particular reason to hold interviews in any particular sequence. Ultimately a scheduled staff meeting meant that the pupil groups had been held first. The pupil groups were conducted on subsequent mornings as pupils were available and returned consent forms. The staff meeting was held a few days later.

3.7 – Research methods

3.7.1 Changes to pedagogy and the intervention tool

The literature review heavily influenced the design of an effective intervention tool. Nation's (2007) four strands concept led to construction of an intervention that included reading, writing, speaking and listening. This was achieved by utilising Mason, Mason and Quayle's (1992) workbook-style approach but simplified into individual worksheets that corresponded to topics. The methodological approaches of August *et al.* (2009) and Lara-Alecio *et al.* (2012) were also a strong developmental influence.

3.7.1.1 Selection of keywords for each worksheet

The course was taught in topics which matched sections of Levesley's (2011) textbook for the course. When considered alongside the department schemes of work 22 main topics were identified to spread the keyword sheets between. Some additional topics existed that did not link to the selected list of keywords (in Appendix A). Most sheets contained around four keywords to focus upon but there were exceptions for topics that required a greater number of keywords to be introduced and used in one lesson, for example life cycles of stars. This did highlight that some topics involved a greater number of keywords that might have previously been consciously considered and future consideration could be given to whether topics should be broken up further to aid vocabulary development. This however was not a

principle consideration of this research and would have been difficult to achieve given the limited number of lessons available.

3.7.1.2 Worksheet format

The worksheets were developed to contain four sections. The first section provided the keywords in a table with space for pupils to record the definitions. This was designated as a listening input exercise. This was followed by a short passage which explained the meanings of the four words. Pupils were asked to use this to check their definitions in the first section. This was a reading input activity. A short speaking activity was then provided to encourage pupils to use speaking output and also listening input. This was tightly focussed around the keywords but attempted to add some variety to the use of the words. Finally a written exercise was given which involved a short task utilising the words within a scenario so as to encourage written output with further variety and a sense of higher level use based around Bloom's (1956) ideas. An extension task was provided but more to ensure pupils continued learning if they finished more quickly. This task was not considered essential to the input/output theory as the "four strands" of word development had already been encouraged. An example worksheet is shown in Appendix D.

3.7.1.3 Use of the worksheet in lessons

The sheets were found to take between 15 and 20 minutes to complete when piloted with the Triple science group. The worksheets also proved accessible to the pupils during this pilot and within this time frame. The boxes in which to write definitions

were perhaps a little small but in order to preserve a readable font size and keep the sheets to one side (which seemed more practicable) there was no easy way to enlarge these boxes so they remained as they were but pupils were encouraged to write outside of the boxes if they needed to. Completing all four tasks proved slower than expected and this immediately highlighted that the benefits of using such a long worksheet would need to be clear in order to justify giving a substantial percentage of a one hour lesson to their completion.

3.7.1.4 Staff CPD for use of the worksheet

Staff required training in the use of this tool. It was important that they understood the concepts outlined here and used them in the appropriate way. Training was provided in a one hour CPD session, followed by more brief discussions on progress in further meetings that occurred during the research. Teachers of the classes involved in the research were then asked to teach using their own planning and approaches within the bounds of schemes of work and school policies, with the request that they used the tool during the lessons. They were allowed to adapt the use of the tool if they felt that this was required. This was in part to ensure pupil examination and learning, as well as teacher responsibility for their own class, was not compromised which could be considered unethical. It was suggested that the use of the worksheets could be in keeping with Bloom's taxonomy (1956) by using the intervention as a "naming" level of work prior to higher level activities.

3.7.2 Quasi-experiment data collection

In order to determine pupil understanding of the keywords before the change in practice and then assess their developed understanding afterwards, a standardised test was used. Borg and Gall (1989, p.292-3) stressed the importance of thoroughly researching tests and in this instance the research had not revealed any test which would be usable for the purposes of this quasi-experiment, at least not in an “off-the-peg” manner. Therefore the test used needed to be created solely for the purpose of this investigation. This removed any ability to make wider comparisons or utilise the literature of standardised tests but did not prevent an analysis of changes in pupil scores between pre and post-test. Given the variables that would differ between schools it would in any case have been difficult to make useful comparison with standardised test scores from other schools. The purpose here was also to establish impact in this setting not at this stage to establish potential wider impact.

3.7.2.1 Testing understanding of the keywords

A three tier approach to testing as identified in the literature review (Waring and Takagi, 2003), using different test types, seemed unnecessary given that the results showed that their three tiers were indicative of each other. As such a multiple-choice format was chosen as being the simplest for pupils to approach and relatively straight forward to mark and score. This would ensure a standardised approach to marking and allow more words to be tested than if a written definition had been required. However, it meant that pupils could guess at words that they were not completely comfortable defining. Conversely it meant that pupils not selecting a

correct answer were unsure even when the correct definition had been presented as one of the options. This would be useful in confirming that pupils did not know the word rather than had been unable or unwilling to try to define it in their own words.

This approach did contain a drawback as pupils selecting the appropriate definition were not necessarily demonstrating their capability to define a word if it had been presented without any form of support or guidance towards the correct meaning.

The format of GCSE papers at this time means that pupils would be able to tackle the common multiple-choice approach often used on the papers but might struggle to use the words in questions building upon the definitions. Many GCSE paper questions are currently “ramped”, meaning that they build upon the multiple choice definition but that does not preclude the possibility of the need to be able to define and use terminology of words not previously presented in the early multiple choice component of the question. It also fails to consider the need for pupils to use the words more confidently in later science studies and in employment. Nonetheless, the study by Waring and Takagi (2003) found that asking pupils to define words in a written form with no prompts produced similar results to the multiple-choice format therefore the results were considered indicative of pupil ability. This was a necessary compromise given the constraints of time precluding a test requiring full definitions of many key words as well as the potential difficulties in doing this for less able pupils. Experience suggested that some pupils would not have engaged with this testing approach. In essence validity is compromised to an extent in order to improve reliability of the testing across a wide range of pupil abilities.

The 100 words chosen for development in the physics module for this research were too numerous to all be included in a meaningful test. The time taken for the test needed to not impact too heavily upon teaching time and ensure pupils coped with the test. This left the question of which words to use. There was no clear justification for selecting words, although length could have been considered, following Baddeley, Thompson and Buchanan's (1975, p.575) finding that memory span was inversely related to word length, as discussed in the literature review. The results vindicated an approach which randomly chose words to test, given no link was found between length and word difficulty for the pupils. There was no precedent for the exact number of words to use but an assumption that it would take around 30 seconds for each word in Section B would result in 50 words taking around 25 minutes to complete was found to be accurate following piloting with the Triple science group.

As in Waring and Takai's (2003) test, giving four options to choose from for the answer seemed reasonable given that this would be diverse enough to ensure probability of a correct answer through guesswork (0.25) would be low but potential answers were not so numerous as to cause excessive confusion or duration of the test. This was the approach also used in the GCSE examination papers by EdExcel. More options would also have been difficult to fit on a page and as such would start to become confusing in terms of layout.

A strict time limit seemed less important than ensuring completion so that understanding was assessed rather than speed of recall. This would clearly not be the same in the examination when working to a time limit would be important but

time management would be addressed when revising at a later point. As such staff were instructed to allow pupils a comfortable time to indicate that they had finished the test.

As the test aimed to assess whether a pupil understood a word it was important that the incorrect options in the multiple-choice test were not only incorrect but also were not closely enough related to result in misinterpretation. However, common misconceptions were included where possible given their importance in ensuring true understanding of definitions in science. Other possible answers were sufficiently different from the correct answer so as not to cause confusion but were related to the concepts surrounding the test word as opposed to being words that were easy to discount because they were obviously unrelated. An “I don’t know” option was also provided to encourage pupils to be honest and reflect that they did not know an answer, rather than make a lucky guess, which would have potentially resulted in less clarity in the results. See Appendix E for a copy of this test paper.

The test was administered by class teachers who were accustomed to setting exam conditions in a classroom. This would ensure pupils were comfortable but could not copy. Exam conditions meant that pupils were not allowed to communicate in the test, teachers could not offer help and no resources were be allowed to be used. Pupils were also sat in a way so as to reduce the likelihood of copying. This varied according to the room used but typically meant no two pupils would sit directly next to each other facing the same direction. There was no pressure placed upon staff to be concerned with test scores and so no reason to be concerned with that the tests were in any way affected by the teacher being responsible for their completion.

The word test was piloted with the Triple set as they began the physics 1 module earlier than the other sets (owing to different structuring of their course) and this pilot indicated that the test was suitable to use going forward and in keeping with expectations.

3.7.2.2 Recording of test data

The test was marked by the researcher (to ensure consistency) with a mark issued for identifying a definition of a word correctly. Multiple selections were not counted but were coded accordingly. Not selecting an answer was also recorded. Results were recorded by pupil (anonymously for reporting in this research) and by word in a spreadsheet to enable thorough analysis of the different aspects of the test. The act of the researcher marking the tests meant that teachers were not able to evaluate pupil success directly and as such they were provided with the full results on spreadsheets so that they could carry out analysis of the results for use during revision. Pupils were also given a list of words answered incorrectly so that they could carry out further targeted revision.

3.7.3 Processing of data from the quasi-experiment

Hinton (1995, p.290) outlined the importance of understanding statistics and not simply entering them into a computer. He went on to provide four questions that should be asked about data after input as summarised below:

1. Look for anomalies that may be caused by data input.

2. Check for missing values that should not be there.
3. Check that the data pattern looks correct.
4. Check that the data is entered in the correct order, particularly for analyses such as ANOVA.

(Hinton, 1995, p.291)

This process was adhered to and resolved minor procedural issues such as variance in names of pupils on test sheets. It also allowed a period of familiarisation with the data and its potential manipulation prior to more extensive analysis.

Rose and Sullivan (1996, p. 46-49) discussed the importance of coding data, emphasising the importance of "...both theoretical concerns...and methodological ones...". Raw data for this study was initially encoded with a number according to which answer had been given by the student for each word. There were only two apparent points of difficulty in coding using this approach – how to code a null response (no selection made including the "I don't know" option) and how to record multiple selections (which were not permitted). Ultimately a lack of response or unclear single response was encoded separately, as were multiple selections (but not so as to identify the choices made in the multiple selection. Since no response would suggest a pupil did not know the answer (though they might have accidentally missed the question or forgotten to come back to it) this coding was not of great overall concern, given the emphasis on correct answers. The raw data was further processed to provide a simpler "Y" (yes) or "N" (no) as to whether a pupil had correctly identified the meaning of each word. This process did mean that it was not possible to subsequently identify the choices made when more than one answer was

selected and this could have been useful in a larger study where patterns might have emerged to indicate whether this was accidental or a point of confusion between definitions. Occurrences were too low in this study to have any statistical significance. An additional consideration in coding was that it was not viable to measure or record whether a student did not genuinely know the answer as opposed to the student becoming disengaged and simply circling “I don’t know” to avoid fully participating in the test. This meant that there was a potential level of error in the data and this error was very difficult to isolate. Therefore the results would have to take account of a level of uncertainty in this regard.

3.8 Survey data collection

Interviews were deemed the most appropriate approach to surveying pupils and staff and as such literature was reviewed further in order to develop an appropriate methodology.

Many approaches to conducting interviews could have been used, from individual structured interview to a loose format approaching the ethnography end of the spectrum (Powney and Watts, 1987, p.5). McAteer (2013, p.34-35) and Powney and Watts (1987, p.102) felt that personal values and sensitivities were of considerable importance in the process. Clearly the use of perception would need to be evaluated appropriately when drawing conclusions but could be particularly important, especially given the setting of a girls’ school where in this particular instance the girls have been seen to respond well to this sort of approach in the past. McAteer (2013, p.73) expressed the opinion that “Interviews...are particularly useful in helping the

researcher to ‘get inside’ the story...” Freebody (2003, p.133) and McAteer (2013, p.73) felt that semi-structured interviews were particularly useful in this process and suggested the use of guideline questions with flexibility around how these were worded. Powney and Watts agreed (1987, p.12) advising “Whatever the style of research, elegance and good experimental design alone cannot compensate for full description of the researcher’s quest for meaningful verbal relationships and their consequences for action.”

Walford (2001, p.96-7) also discussed the importance of observation of the unspoken in research whilst Powney and Watts (1987, p.144-5) discuss “unrecorded data” and considerations that needed to be given to recorded things that are not said, retaining constantly an awareness of interpretation and the position of the interviewer. In this analysis the “unsaid” and the general observed feelings from the group were noted first whilst the transcripts were edited only to include gestures that indicated agreement where relevant.

3.8.1 Interview methodology

Focus groups were used for staff and pupils. The groups were interviewed using a schedule of questions (discussed in detail in 3.8.1) but subject to the discretion of the researcher where there was overlap or a need to probe further. Guiding questions were however avoided and the researcher’s skills in reflective interviewing were utilised to avoid this. The interviews took place in science labs which proved comfortable and were made free from distraction by conducting the interviews at quieter times in the morning for pupils and after-school for staff. The seating was

also adjusted to facilitate eye-contact between interviewer and interviewees. This responded to concerns over the setting and seating arrangements by McAteer (2013, p.76) and Lewis (1992, p.418) respectively.

The focus groups were recorded (using two recorders, one as a back-up) with minimal note taking of non-verbal interaction and allowing a more comfortable interaction with the participants. The records were transcribed as soon as possible after the focus groups (within 24 hours to maximise recent recall of events).

As previously discussed in the research design, interviewing was deemed an appropriate method for conducting a survey (Powney and Watts, 1987, p.4). Corte, Verschaffel and Van de Ven (2001) had also successfully used interviews as the key method in their evaluative approach. For the purpose of this thesis it seemed appropriate to develop the interviewing methods that would develop feedback akin to a first review during action-research, following the approach of Corte, Verschaffel and Van de Ven (2001). This meant discounting Sapsford's (1999, p.119-135) approach involving questions which were more reminiscent of a questionnaire. This allowed for more flexibility in the questioning (as discussed in the method) in order to maximise the strength of this design frame and methodology in giving the interviewer the ability to probe into areas of importance as they arose. This was useful several times given that pupils and staff had plentiful feedback and developmental ideas. It did mean that results were influenced by the researcher-interviewer. However, this did not appear to reduce the value of the final conclusions. Thorough action-research would have been more appropriate for refining the intervention but survey was suitable for discovering whether there was

evidence to justify a longer study. Ultimately the study proved that further refinements were possible before investing in a longer study which would suit more precise refinement.

It should be noted that Walford (2001, p.88) felt there were problems with the interview approach, saying "...every person who is interviewed carries his or her own construction of what 'an interview' actually is..." and "...few of these conceptualizations coincide with the relationship that most qualitative researchers would wish to establish." The important point seems to be ensuring that the method was suited to the objectives and that the data gathered could potentially be used for analysing the answer to the research questions posed. This was the case so Walford's fears (2001) were allayed.

Walford (2001, p.92-3) also noted that during his research, including that on physics research students, that he had wanted to observe what was going on for several weeks before conducting interviews. He went on to say (Walford, 2001, p.93) that this "...was simply not possible" due to him being considered a "questioner" to those he wishes to observe. However, this study found that within a school it was possible to act as both researcher and interviewer to good effect, noting that this research was relatively small in size and focussed in one school with the researcher a key link which might not be replicated easily in other settings.

3.8.2 Nature of the interview and schedule of questions

Following the recommendations of McAteer (2013) and Freebody (2003), semi-structured interviews were used in keeping with the points in the previous section. The schedule of questions needed to elicit the perceptions of pupils and staff with relation to the pedagogy and this was to be of key importance in the success of the interviews (Lewis, 1992, p.420). Avoiding leading questions was important but it was useful to question pupils and staff more closely on arising points as the interview proceeded. McAteer (2013) and Powney and Watts (1987) felt that personal values and sensitivities were important in interviews and as such the questions were worded so as to elicit “feelings”. Powney and Watts (1987, p.100) expressed a preference for group interview to “elicit from students their retrospective views of, and current feelings about, their experiences of science education 11-16.” This had proven a successful approach with the pupils in the past in this school setting too. With this in mind the schedule of questions for pupils was as follows:

- How do you feel about your science this year and how well you are making progress?
- What difficulties do you find that you have in science?
- What do you feel you do well at in science?
- What activities in class do you find most helpful?
- Do you find that using the keyword sheets is useful?
- Do you feel these are these more useful than your own revision, such as using flashcards or revision notes?
- What do you feel would help you make better progress in lessons?

Whilst the teaching staff were not considered to require the same use of emotive language it did seem appropriate to gauge opinion in keeping with the questions asked to the pupils. The use of standardised assessment of lessons and teaching was also of concern since this was not a focus of this research. As such the continued use of feeling in the questions was aimed at allowing staff to express how they felt about issues as opposed to how they might have been expected to feel. As such the schedule of questions for staff was:

- How do you feel about your teaching for the GCSE classes this year?
- What do you feel is going well with the GCSE classes this year?
- What do you feel are the major difficulties in teaching pupils the GCSE course this year?
- Have you had any success with particular activities in GCSE lessons this year?
- How do you feel about the keyword sheets?
- Do you feel the keyword sheets are a useful use of time compared to other tools?
- What do you feel would help improve lessons further (with regard to the keyword sheets or just in general)?

These questions were deemed as being suitable for gathering the data which was used alongside the quantitative analysis to answer the latter two research questions, which were:

- What worthwhile benefits do pupils perceive from the change in pedagogy?

- What worthwhile benefits do teachers of the pupils perceive from the change to pedagogy?

3.8.3 The role of the interviewer

Powney and Watts (1987, p.7) said that there were frequently issues with untrained interviewers with "...limited interviewing, educational or other relevant experience."

They (1987, p.34) went on to point out that "In small-scale educational research, researchers often carry out their own interviews." They did also highlight issues with experience and commitment to the outcome of the project. They noted (Powney and Watts, 1987, p.40) that "Interviewees need to trust the person interviewing them."

The joint role of researcher and interviewer as well as teacher ought to allay some of the concerns about the importance of interviewees trusting the interviewer (Ely et al, 1991, p.93-99). This supported the decision for the researcher to act as interviewer with the level of trust with pupils combined with the need to manage resources for the small project being allayed to some experience in interviewing, though noting that a more experienced academic researcher could have provided additional skills and potentially uncovered additional detail, though there was no apparent evidence for any shortcomings in the interviews.

3.8.4 Considerations surrounding Interviewees

Powney and Watts (1987, p.48) had concerns with child interviewees and pointed out that "They have spent all their years since birth coming to grips with parental demands and all their school lives working out what teachers want and how to

please them.” This was a consideration that was carried through to the analysis stage, though the Year 10 age group (of fourteen to fifteen year olds) and familiarity of pupils with focus groups meant that these concerns appeared unfounded – pupils were confident and constructive in their answers to the questions and in further reflection. This indicated the very different culture in modern schools from that described by Watts and Ebbutt (1988, p.211-219) who felt at the time that pupils were rarely given an opportunity to discuss their science teaching. This stark contrast added to the need to reconsider older research in light of modern practise – the students were previously exposed to many forms of research designed to gather their opinions on a variety of topics, such as curriculum and learning experiences.

3.8.5 Validity

In keeping with the concerns of McAteer (2013), there was time left to reflect between the first and second pupil focus groups in order to reflect on the interview. This time was relatively short being 24 hours but no alterations to the process were deemed necessary. Wellington’s (2000) recommendation for a moderator to be involved was not upheld as it was deemed (and found to be) unnecessary and potentially costly. There was no evidence of this proving to cause issue with the data or conclusions. Sikes’ (2000) previously discussed concerns about interviewees not telling the truth was also found not to be of concern in this study.

Walford (2001, p.88) pointed out that interviews were “...not a transitory conversation, but one that is invested with significance.” He said that “...every person who is interviewed carries his or her construction of what ‘an interview’

actually is.” He also suggested that “The word ‘interview’ might not be used at all, and the process might be seen as a ‘guided conversation’, but people are not fooled.” In this research the term “focus group” was adhered to since pupils were familiar and comfortable with this term – it had been used for routine surveying of opinions without problem in the past.

3.8.6 Piloting of the focus groups

The first pupil focus group was effectively the pilot, meaning that it followed the ethics procedure outlined. There were no problems with the timing of the group or with the proposed questions and pupils appeared comfortable and able to participate fully. The only modification was that a considered use of more comfortable office space proved small and impractical to use at the required times (with staff also needing access to the space). A lab was used to conduct the groups and this was in keeping with normal school activities with pupils. The desks were movable so as to create a small group work area and the pupils did not appear to react to this in any way which would be adverse or with impact on the results.

3.8.7 Analysis of the Data

3.8.7.1 Introduction to the analysis of focus groups

This sub-section addressed the way in which the interviews were analysed, cumulating in transcripts used to answer the research questions and thematic

analysis identifying additional ideas that supported the wider evaluations of the research.

3.8.7.2 Supplementary information about the focus groups participants

In summary, two pupil focus groups were conducted. The first consisted of five pupils and the second consisted of six pupils, in keeping with the research design. Pupils were asked to participate through discussion and an accompanying letter, as described previously and included in Appendix C. Pupils were sampled by the top, middle (mean) and lower end of each set using the post-test scores from the word test. All of the pupils approached indicated that they were happy to participate however the ethical requirement of obtaining parental permission was a problem and a small number of pupils failed to obtain consent, even with polite reminders over three mornings subsequent to the initial discussion. The language spoken at home was one issue encountered but one which could not be resolved in time for the focus groups to take place soon after the completion of the use of the word sheets and tests. Given the connection between language in this thesis and the particular issues surrounding pupils who speak another language at home this is was an important point within evaluation of the research. A truly representative investigation needed to better tackle the issue of EAL pupils, with an interpreter required. In some instances a sibling could have facilitated this but this was not possible in this study and it would have been necessary to engage a translator, be that a professional or a suitable bilingual adult who was willing to help contact home or translate letters. The facilitation of this in the time frame available was not possible, especially when cost considerations (which needed to remain minimal) were also taken into account.

One pupil was substituted by another with a similar mark whilst a further three did not return consent forms and did not participate (nor were they substituted due to time constraints). Nonetheless the participating students represented all five sets and the range of final scores. Those pupils not happily engaging with the final test were not approached and again it is important to be aware that the focus groups are not representative of pupils known to have a particularly poor attitude at this stage (it was deemed unlikely that they would agree and/or consent to be involved productively). This again meant that the research presented an issue with a truly representative study.

For the teacher focus group six out of seven staff teaching the course attended and consented to participate in the staff focus group. The group covered the three subjects specialisms, part and full-time staff and permanent and non-permanent staff. The six staff represented teaching of all five sets. One staff member was unavailable due to other commitments but had been fully supportive of the research activity.

3.8.7.3 Analysis of the “unspoken” aspects of the focus groups

In response to Walford (2001) and Powney and Watts (1987) the unspoken elements of the focus groups were considered as a part of the analysis. The first pupil focus group was made up of those five pupils who returned their consent forms first. There might have been some significance in this selection for the first group. The pupils were clearly keen, conscientious and/or organised in that they spoke to their parents

very quickly in order to secure their consent. The parents may have had an impact on these factors themselves, perhaps helping to organise the pupils. This highlighted further factors that might have been at work within the full understanding of pupils approach to science and therefore their outcomes in the subject as well. Four of the five pupils were particularly positive in their responses and the group gave the impression of being largely happy with their progress, the intervention and their experiences in studying GCSE Science generally. They were quite willing to discuss their experiences confidently. Three of the pupils were in the Triple set taught by the researcher and as such there might have been some influence over those pupils by their involvement with the intervention in a more direct sense than the other sets. The pupils were forthcoming with suggestions for improving their learning further.

The second pupil focus group had a very different feel. The pupils were less familiar to the researcher/interviewer – possibly highlighting the concerns that can arise from the researcher also acting as interviewer (Powney and Watts, 1987, p.118). However, Walford (2001, p.95-7) talked of his positive research experiences in being the interviewer in various situations though noting the limitations of the research frame. This group of pupils came across (with one exception) as much less happy about their progress in science. They seemed much less confident in the subject. To some extent this was impacted by the pupils opting to study science further (through the option of Triple science) being in the first focus group. The pupils in the second group had opted not to expand their science studies for their GCSE studies and as such may have been less interested in the subject or may at least have less confidence relatively. However, the original motivation behind this study was to help

all of the pupils with their science studies and the keywords approach was required to help those pupils who were least secure with the subject as much as any other pupil. The pupils did interact well with the questions and as such they provided a full set of responses – although it was also apparent that some of these pupils had less familiarity with the keyword sheets than others.

The staff focus group was supported in a positive way by the staff. They all contributed, bringing a balanced but mixed set of opinions and viewpoints on the keywords and progress with the groups in general. There was an overall positive feeling to some of the ideas and some clear shared ideas about how to utilise the learning and how to compliment it. It was clear that staff had interacted with the available resources to differing extents.

3.8.7.4 Using transcripts to address the research questions

Powney and Watts said that “Part of the creative process in analysis is to impose a structure on the accumulated material” (1987, p.11). Freebody (2003, p.133) said that following interviews “...talk is typically then tabulated or transcribed in full then the researcher may decide what to analyse in depth, depending on the patterns and themes that emerge.” However, Ely (1991, p.91) pointed out that “Text analysis packages do not do the analysis for the researcher. The user must still create the categories, do the segmenting and coding, and decide what to retrieve and collate.” Ely (1991, p.91) also warned that the computerised systems have been found to put a gap between the researcher and research. The use of NVIVO to store and organise data did not appear to create this issue, perhaps due to the researcher’s

familiarity and comfort with computerised systems. The software was useful for reviewing the transcripts and organising the ideas.

Initially the interviews were transcribed to provide a complete record of the conversations. Following the suggestions of Powney and Watts (1987) and Ely (1991) discussed in the literature review, data from the interviews was organised into themes in two different ways. First of all the transcriptions were analysed for comments which directly responded to the research questions. This did not require more complex reorganisation of the data since the interview questions were designed to relate to these research questions. However, it was clear that both pupils and staff had discussed ideas which went beyond straight forward answers to the research questions and as such a secondary analysis of the transcripts was required in order to organise these ideas into themes and explore their relationship to the research.

3.8.7.5 Thematic analysis of the interviews

Ely (1991, p. 143-4) suggest that “thinking units” could be used as a starting point for analysis of interviews. Following these ideas an analysis of the data was carried out using thinking units which were then developed into themes. These themes were used to ensure that useful ideas that developed from the interviews were considered either within the separate analysis for each research question or as part of the further findings in Chapter 5. The thematic element of this approach mirrored the successful method utilised by Francis, Hutchings and Read (2004, p.6) who also

used interviews to develop findings, although from questionnaires rather than a measured intervention.

Initially the interviews were transcribed in full, verbatim, after being conducted. The number and length of the interviews meant that there was no reason to attempt to abridge this process. This also allowed for revisiting of comments as further ideas developed during the analysis. A process of thematic analysis using coding of comments (within NVIVO) led to the development of key themes for pupil and staff focus groups separately. Initially the themes arose from two branches – the word sheets or general points related to learning in science lessons. A review of the end result led to the development of a different set of themes based around prevalent ideas across the interviews, for example “key words” and “prior learning”. These themes were used as a reference point when reflecting upon the developing analysis and also to ensure that the data from the interviews was fully examined and included across the final analysis.

Thus Chapter 4 incorporates interview data where it supported deeper analysis of the first research question where the statistical data pointed to conclusions which required deeper understanding. Chapter 5 considered the second and third research questions primarily using the interview data as the more prevalent sources. The additional themes were then analysed in Chapter 6 where they provided a wider understanding of the research findings.

Chapter 4 – Analysis of Research Question 1

4.1 Introduction to the analysis of findings

The findings of this research were addressed in order of the three main research questions. However the quasi-experiment results were processed and analysed alone first of all in order to fully establish what the quantitative data had revealed. This data was the prevalent source for answering the first research question: Do pupils show an improved ability to define keywords following the introduction of a pedagogical intervention in a GCSE physics module? Data from the focus groups were also included where it was helpful to gain a further understanding of the quantitative results directly or more indirectly where deeper issues emerged. The analysis was further sub-divided into consideration of the individual pupil scores, the change in scores for the individual words and possible explanations for the variance in scores, including issues around EAL pupils and the development of the word test.

4.2 Analysis of data for all pupils by pupil score

Pupil scores correspond to the number of words a pupil correctly defined in the test. This was analysed using descriptive statistics, as shown in Table 4.1 below.

Table 4.1 – Descriptive statistics for pupil scores pre-test and post-test

	Pre-test	Post-test
Mean	25.3	35.2
S.D. (std.)	8.7	9.7
Max score	45	48
Min score	3	4

Before the intervention (and the teaching of the corresponding GCSE topics in Year 10) the pupils gained a mean score of 25.3 words. This meant that on average, pupils were only able to define about half of the keywords used in the test. As concept words and more complex vocabulary was selected for the intervention and test it would be difficult to accurately extrapolate this result to all keywords in the topic but it did suggest that given the random selection of 50 words from an original 100 that pupils would find use of at least 50 words in this topic difficult after being taught them. The standard deviation of 8.7 indicated that there was considerable variation between pupils. The maximum and minimum scores indicated this further, showing that pupils ranged from knowing only three words through to forty-five. This indicated that the need for an intervention on keywords was significant for some pupils but almost unnecessary for others, even prior to teaching of the module. This in turn suggested significant variance in prior learning connected to the vocabulary either through teaching in school or learning away from school. The later analysis of word familiarity began to address this complex link.

After the intervention the mean increased to 35.2, showing that pupils had improved their ability to define keywords by around ten words on average. However, the mean was still significantly below the possible maximum of 50 and so it had not ensured that the words were understood by all pupils. The maximum score had increased by three to 48, demonstrating that no pupil was able to define all of the words even after the intervention but that most (but not all) pupils had improved to some extent. The minimum raised by only one, simultaneously showing that the intervention was not successful for some pupils. This matched comments about the word sheet intervention which varied from "...the sheets do help..." to "...I can't learn from the sheets..." Staff had picked up on this too saying that "some of them clearly don't know how to revise." Pupil development of vocabulary was very varied.

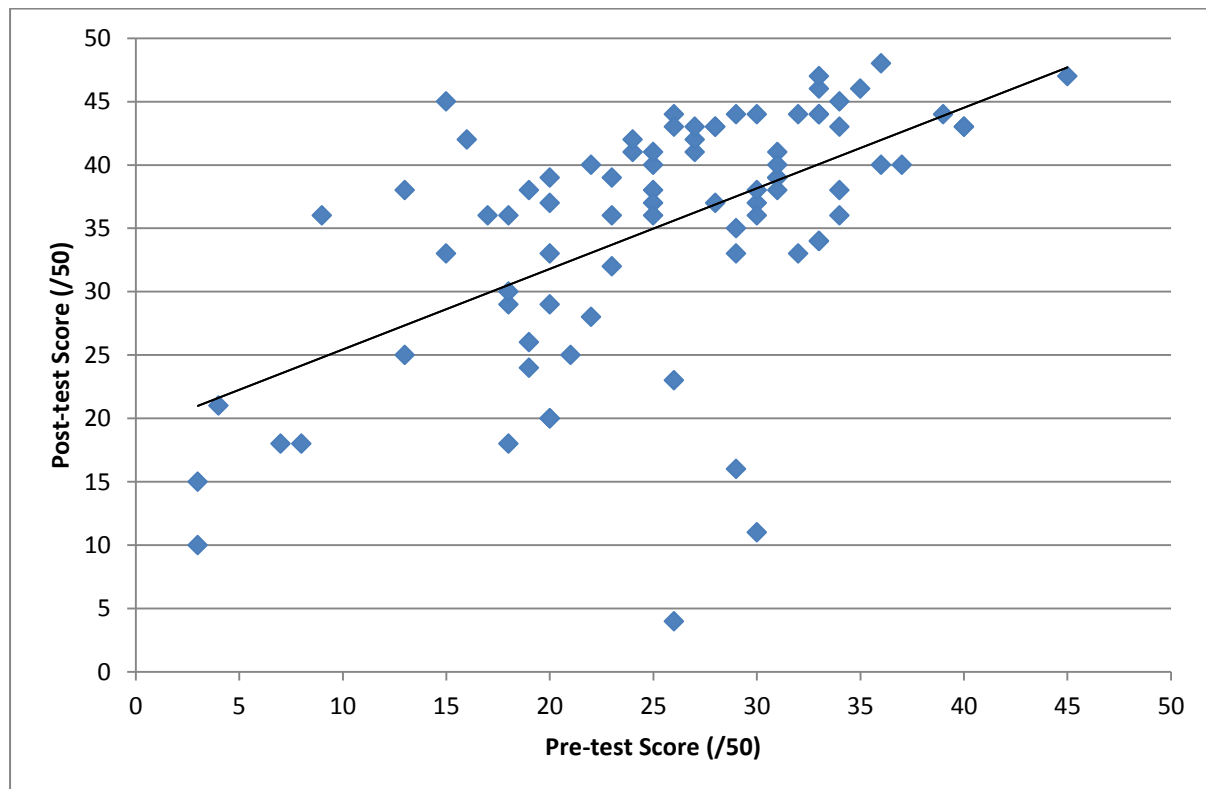
Further analysis of this data showed that two pupils did not improve their score at all. A further four pupils had a lower score after the intervention. Concerns raised by staff about a small number of pupils not approaching the post-test with a positive attitude were believed to contribute to this. This did link to the concern about the reliability of the test but was useful as it showed that some pupils were responding negatively after the teaching of the module and that the success of the intervention was linked to pupil attitude, seemingly conflicting with the findings of Galton, Gray and Ruddock (2003) that attitude was not necessarily a link to outcomes in science. More in-depth individual studies of those pupils could have revealed links to motivations and aspirations which would have linked to the wider findings of the literature review, particularly the research by Gorard, See and Davies (2011) and Goodman and Gregg (Eds., 2010) but future design would need to consider how

best to engage with these groups of pupils successfully – possibly using an evaluative method other than interviews.

The greatest increase in score was a value of 27. This was from a pre-test score of nine and showed that the teaching and intervention during the module, combined with the pupils' own additional studies, had results in a significant increase. Again this raised further questions that required the survey to begin to analyse. For example, one pupils indicated that she used the sheets to revise regularly, saying "...if I'm doing a test I might look over the keywords." Another pupil said "...they're good...for a quick recap..." Whereas another pupil said "...I can't learn from the sheets..." and others indicated varying use of the sheets either due to teaching variation or their own effort to complete tasks fully. These comments showed that one approach was not suitable for every pupil but the early concerns about pupil aspirations and motivations also appeared to be at play. One pupil highlighted the latter point in saying "...I hardly ever used to answer the questions..." with another saying that there was a need for teachers to ensure everyone was focussed on the tasks. This showed an awareness of the different levels of motivations to complete the work even amongst the pupils.

So overall pupil scores had improved but with varying degrees of success. A comparison of all of the pupils' score changes was required to better understand the data. As such the individual results were also plotted as a scatter-graph to show the trend in the relationship between pre-test and post-test scores, as shown in Figure 4.1 below.

Figure 4.1 – A scatter-graph showing the correlation between pre-test and post-test scores for each pupil



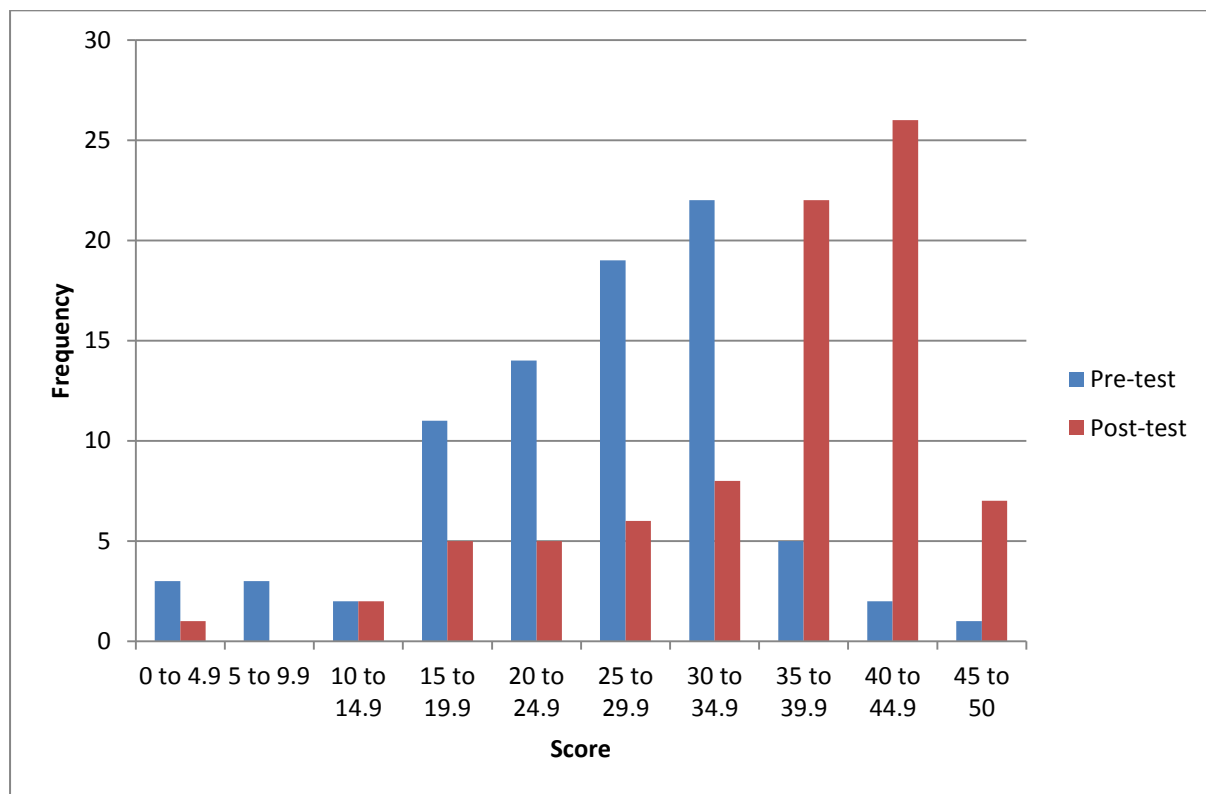
The graph showed a positive trend, indicating that as a general trend pupils had improved their scores. However the improvements were smaller as the pre-test score becomes higher. As no pupils gained the maximum score there was no reason to believe that the fifty word limit had caused this limitation and instead suggested that achieving the maximum score was difficult. In turn the logical assumption here was that certain words were remaining difficult for pupils to define (at least using this test format) and this was explored in the analysis of the individual word scores in the next section. It was however possible that there would always be some level of error for pupils owing to slight confusion or misplaced selections,

although this ought to still apply to lower scoring pupils in proportion, countering this idea.

The graph also clearly showed outliers, particularly in the 25 to 30 pre-test score range. They corresponded to the pupils with lower marks and the size of the difference again supported the idea that the cause was not a simple loss of understanding and that there was a flaw in the use of testing to gather data. Staff indicated that they felt some pupils had deliberately not attempted to complete the second test to the best of their ability. This raised issues with the reliability of the test (though clearly some of the anomalous results could be identified and processed). Tapping into the issues with disaffected pupils was difficult despite the importance of understanding the attitudes, aspirations and motivations of these pupils. There were also points that were well above the normal trend in the 10 to 20 pre-test score range as well which suggested the possibility that pupils having some pre-existing knowledge but finding the definitions difficult had particular potential to improve their score with the intervention. This seemed to be linked at least in part to those pupils engaging with the process and revising from the sheets further. This was again supported by the aforementioned comments demonstrating the variance in approach by pupils towards the sheets in terms of their own revision. Some pupils clearly found the sheets useful whereas others did not.

A clearer understanding of the shift in the distribution was seen when the data was displayed as a frequency graph, as shown below in Figure 4.2.

Figure 4.2 – A frequency histogram for pupil score pre-test and post-test



There appeared to be an overall shift from a normal distribution centred mid-way in the range of possible marks through to a pattern shifted towards the higher scores and bunching around the highest score categories. However this was not certain given the relatively small amount of data, which led to few categories and the potential for high impact of anomalies. A statistical calculation of the change was made to investigate the possible shift in the pattern further. A “paired-samples t-test” was used (Thomas, 2009, p.224) based on the null hypothesis: “There is no link between the use of the intervention alongside teaching of the module and pupil ability to define keywords”.

Thus t was calculated as 90.78 for 81 degrees of freedom. Such a large number had a very high level of significance (easily beyond the standard of 0.05). Thus the null hypothesis was rejected and the hypothesis was accepted, in that there was a positive correlation between the use of the intervention alongside teaching of the module and pupil ability to define keywords. However, the level of impact of the intervention alone was not clear due to the use of the quasi-experiment for the reasons already discussed. A control group would have been needed to do this and it would have needed to be of a size substantial enough to generate a level of statistical significance which was not likely to be achieved using a population of this size.

The results were in keeping with the findings of August *et al.* (2009, p.358) in that they also saw an increase in the score of their vocabulary test post-test (by a factor of 1.6) and an increase in standard deviation as well (by a factor of 1.3). This compared to an increased score for this study of a factor of 1.4 for outright score and 1.1 for standard deviation. It should be noted that August *et al.* recorded their scores separately for ELL and non-ELL pupils – the scores quoted here were from non-ELL (ELL scores showed a smaller improvement). EAL/ELL issues were considered later in this chapter, being exceptions rather than a significant sub-sample in this study.

4.3 Analysis of data by word score

The analysis of data using the pupils' overall scores was useful for considering the variation between pupils and for gaining an understanding of the overall success of teaching in terms of understanding of vocabulary. However, the considerable

variation in individual scores and difference between pre-test and post-test required further analysis of the data to try to find a deeper understanding of the factors at work.

4.3.1 Improvement in score for individual words

Having considered the prior knowledge of words by pupils it was necessary to consider the pupils' improvements in terms of the different words. The descriptive statistics for the scores by word (rather than pupil) pre-test and post-test are shown below in Table 4.2 for comparison.

Table 4.2 – Descriptive statistics for success in terms of individual words

	Pre-test	Post-test
Mean	50.6%	70.3%
S.D. (pop)	24.1%	17.8%
Max score	96.3%	96.3%
Min score	11.0%	24.4%

The descriptive statistics showed that the number of correctly defined words rose to 70.3% from 50.6%. The standard deviation dropped from 24.1% to 17.8%. This meant that words had been defined correctly by more pupils on average, though this was not a new discovery since these descriptive results are tied into the pupil score results. The decrease in standard deviation did suggest that pupils were gaining

more similar results when the individual words were analysed, further suggesting that whilst pupils' ability to define words was varied there had been a decrease in the variation by word meaning that consideration of whether particular words were causing difficulty was necessary. This was supported by pupil comments that included "...sometimes I'm not sure of the meaning..." and "...you've just got the word on its own and you really need to be able to link it to other things to be honest." Here the pupils had supported the concerns of Svensson et al. (2009) that pupils often did not acquire the correct meanings of words. This being the case further development of words would be needed in context with word sheets or even glossaries only acting as a reminder. One pupil suggested this was how she worked with the intervention, saying "I think they're good for sort of like a quick recap if you've got a reasonable understanding of what the word would mean."

The highest and lowest score gave some further insight as to whether particular words were causing more difficulty. The highest score of 96.3% was unchanged between pre-test and post-test and did suggest that the nature of the test meant a level of error would be likely, preventing an increase to 100%. However, the aforementioned difficulties with a small number of pupils might have also been responsible and the 3.7% failure might have been due to pupils deliberately not trying to answer the word correctly. Further review of the results showed that in many instances the pupils with the negative scores had answered these high scoring answers incorrectly, affirming that the lack of 100% scores was due largely to pupils effectively sabotaging their result. Thus the deviation from 100% in these words was not deemed of importance for further understanding beyond this observation –

though a few pupils who appeared to comply with the testing did get answers wrong on these presumed easier words to define.

The minimum score on individual words rose considerably from 11.0% to 24.4%, showing that at least one-in-four pupils (approximately) could define any given word after the teaching of the module. This meant that the teaching and intervention of the module had ensured improvement, including more difficult words with seven words originally scoring less than 24.4%. Nonetheless there were many words which were not being successfully defined by large numbers of pupils despite the specific focus on these words within the intervention and then the teaching of the topic on top. In order to better understand the variation in improvement of each word Table 4.3, shown below, was used.

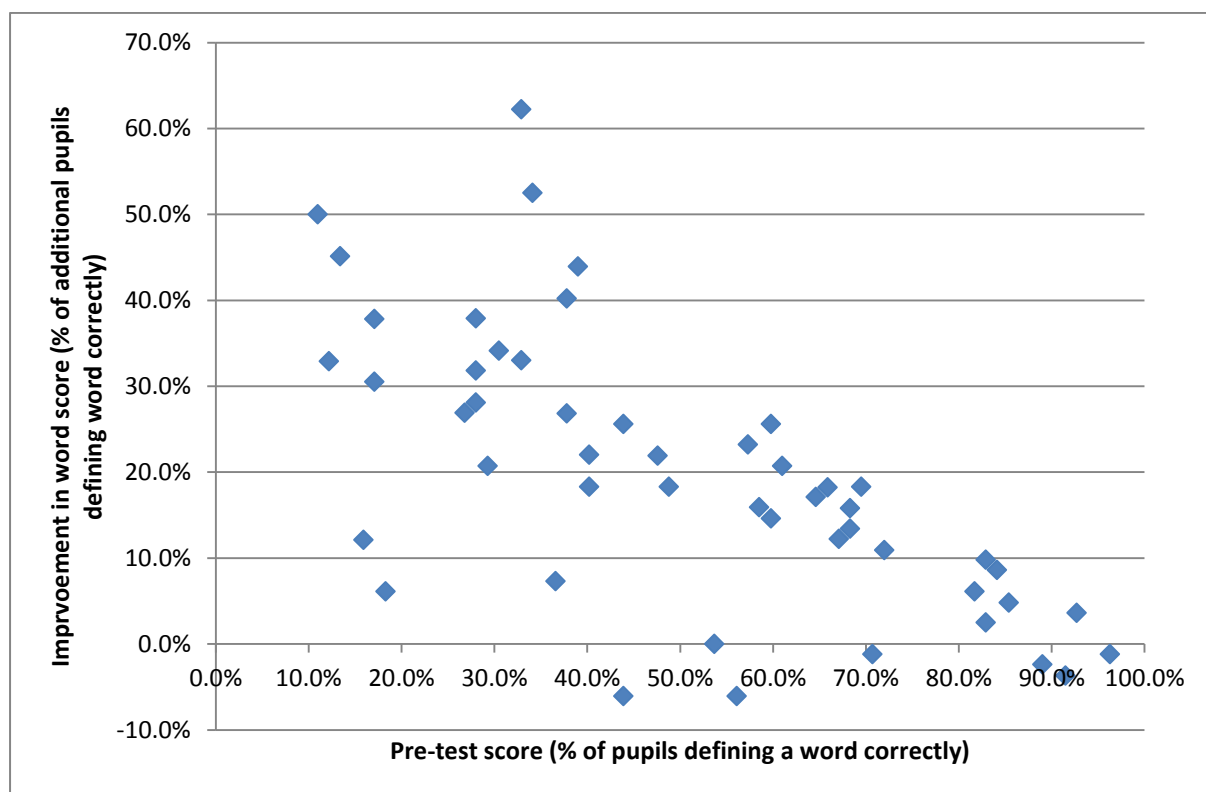
Table 4.3 - Word scores ordered by percentage of pupils defining each word

correctly

Total Correct Answers					Total Correct Answers				
Word	/82	%	Pre-test %	Change %	Word	/82	%	Pre-test %	Change %
Radioactive	79	96.3%	92.7%	3.6%	Electromagnetic radiation	57	69.5%	43.9%	25.6%
Seismic waves	78	95.1%	32.9%	62.2%	System	57	69.5%	47.6%	21.9%
X-rays	78	95.1%	96.3%	-1.2%	Tidal power	57	69.5%	70.7%	-1.2%
Light waves	76	92.7%	82.9%	9.8%	Ultraviolet	55	67.1%	48.8%	18.3%
Thermal imaging	76	92.7%	84.1%	8.6%	Sonar	54	65.9%	32.9%	33.0%
Sound waves	74	90.2%	85.4%	4.8%	Steady State theory	54	65.9%	28.0%	37.9%
Supernova	72	87.8%	69.5%	18.3%	Chemical potential	53	64.6%	37.8%	26.8%
Thermal	72	87.8%	91.5%	-3.7%	Main sequence star	53	64.6%	30.5%	34.1%
Tsunami	72	87.8%	81.7%	6.1%	Elastic potential	51	62.2%	40.2%	22.0%
Real image	71	86.6%	34.1%	52.5%	Infrasound	50	61.0%	11.0%	50.0%
Sound wave	71	86.6%	89.0%	-2.4%	Primary coil	49	59.8%	28.0%	31.8%
Geocentric	70	85.4%	59.8%	25.6%	Protostar	48	58.5%	40.2%	18.3%
Renewable energy resources	70	85.4%	82.9%	2.5%	Slip rings	48	58.5%	13.4%	45.1%
Direct current (DC)	69	84.1%	65.9%	18.2%	Induced current	46	56.1%	28.0%	28.1%
Microwaves	69	84.1%	68.3%	15.8%	P waves	45	54.9%	17.1%	37.8%
Electrical current	68	82.9%	72.0%	10.9%	Acid rain	44	53.7%	53.7%	0.0%
White dwarf	68	82.9%	39.0%	43.9%	Beta (β) particles	44	53.7%	26.8%	26.9%
Alternating current (AC)	67	81.7%	64.6%	17.1%	Hydroelectricity	41	50.0%	56.1%	-6.1%
Tectonic plates	67	81.7%	68.3%	13.4%	Transverse waves	41	50.0%	29.3%	20.7%
Wavelength	67	81.7%	61.0%	20.7%	Neutron star	39	47.6%	17.1%	30.5%
National Grid	66	80.5%	57.3%	23.2%	Step-up transformer	37	45.1%	12.2%	32.9%
Kinetic	65	79.3%	67.1%	12.2%	Converging Lens	36	43.9%	36.6%	7.3%
Ionising radiation	64	78.0%	37.8%	40.2%	Non-renewable resources	31	37.8%	43.9%	-6.1%
Electromagnetic spectrum	61	74.4%	59.8%	14.6%	Focus	23	28.0%	15.9%	12.1%
Infrared	61	74.4%	58.5%	15.9%	Pitch	20	24.4%	18.3%	6.1%
					Mean No. Correct	57.7	70.3%	50.6%	
					Mean Change	16.2	19.7%		

The table emphasises the considerable variation in improvement between words, particularly when regarded in conjunction with a scatter-graph. This scatter-graph gave a much clearer appreciation of the link between pre-score for a word and the improvement in the percentage of pupils defining the word correctly. This clarified whether certain words were inherently difficult and improvement was limited on lower scoring words or, as was the case, if certain words were unfamiliar initially but this had little bearing on pupils' ability to define that word after teaching and the intervention to ensure that they were familiar with the word. The graph is shown in Figure 4.3 below.

Figure 4.3 – A comparison of pre-test word scores to improvement in score



The graph showed a general overall negative correlation, meaning that the lower the initial word score was the greater the improvement in score in the post-test. It was

not surprising that improvement was only small for those words defined well in the pre-test since the high scores could not improve very much in many instances anyway. For example, “x-rays” scored 96.3% in the pre-test so could only have improve by 3.7% at best (to reach 100%). Conversely infrasound scored 11% initially and could have theoretically improved by up to 89%. This meant that this analysis had an unavoidable inherent flaw and this analysis needed to be used with caution. Irrespective of this flaw the outliers in the data were still of interest and are clearly displayed in the graph. The following words had shown less improvement than the general trend: “focus”, “pitch”, “converging lens”, “electromagnetic radiation”, “acid rain”, “hydroelectricity” and “tidal power”. Conversely the following words had shown better improvement than the general trend: “seismic waves”, “real image” and to a lesser extent “white dwarf” and “chemical potential”.

The word “pitch” gave particularly interesting insight since it was given only passing mention in the course text book (Levesley, Ed., 2011), with sound waves not being a particularly significant topic in this module. It seemed conceivable that there was a lack of focus on this word during teaching due to the brevity of the topic. Similarly, the word “focus” was a very precise part of earthquake terminology that could have been easily overlooked compared to “epicentre”. However, this did not appear to explain difficulty with “electromagnetic radiation” or “converging lens” which had been part of substantial topics in the module. It conceivably explained “hydroelectricity” and “tidal power” which although featuring in the intervention would have relatively minor focus due to the size of the topic they sat within. “Acid rain” was linked to this topic so perhaps renewable energy was a particular issue as a topic. It was harder to find any link to the words which had been defined so much

more successfully post-test compared to pre-test. The four words are not linked and do not appear to have any topical link nor a shared category in terms of the category of vocabulary that they belonged to, for example linked to Wellington's (2000b) aforementioned hierarchy of scientific vocabulary. "Real image" was considered to be a conceptual idea and should as such have been more difficult for pupils than a descriptive term like "converging lens". These anomalous words were found to be in conflict with Wellington's general idea of word levels. This did not mean that Wellington was incorrect given the small number of words studied and the use of specific examples here but it did show that the general rule could not be presumed to be correct for all words.

The appearance of the word "focus" as one which the pupils' had difficulty with raised the possibility that words which have multiple meanings were causing difficulty as discussed by Ross, Lakin and Callaghan (2000). Focus would have been met in the lenses and telescope topics in this module too and this might have caused confusion. This might have been a factor in the word pitch, with its multiple meanings, but this was difficult to conceive as an explanation for the other words (though pupils have been noted to confuse hydroelectricity with hydrogen fuel during previous teaching of this topic). The words with the exceptional increases in scores did not appear to be linked to multiple meanings (except for "white dwarf" but this would be a very apparent alternate meaning if applied to its potential other uses), adding some weight to this idea as a partial explanation of the results.

Given this second possible explanation again only seeming to apply in part it seemed conceivable that several factors were at play in explaining the levels of

success with words. If that was the case then it would be difficult to isolate the issues for each word. An extended investigation would need to focus extensively on pupils' difficulties with the words. This would need a thoroughly considered research design which might involve extensive interviewing or a questionnaire. Studying the number of words in this study through a survey in this manner would be very time consuming, involve a change to the ethics procedure and would need to take place with many pupils to establish how consistent the issues detected were between pupils. This would be very useful but was not realistically possible within the confines of this piece of research.

However, further consideration of the prior learning of vocabulary and the effect of this upon this intervention was possible. When considered in conjunction with the overall rise in scores from 45% to 89.9%, the idea that language acquisition is left largely to chance prior to the word intervention was seen to be supported in part. After the intervention the results were starting to form a distribution more in keeping with an examination score (i.e. closer to a normal distribution). This was perhaps not surprising given that pupils have had been specifically taught the words by this stage. This highlighted a possible need to organise the teaching of the GCSE vocabulary prior to the GCSE cause and not simply leave progression of vocabulary to chance or the design of the Key Stage 3 National Curriculum. This supported the ideas on curriculum and learning organisation discussed by Bloom (1956) and Funk (2012). It was also seen to support Francis, Hutching and Read (2004) in their claim that curriculum in science was too content heavy. Analysis of the pre-test scores was used to gain better insight into the importance of these issues.

4.3.2 Analysis of pre-test scores for individual words

Further understanding of why particular words were causing difficulty was gained by analysing the pre-test word scores. The pre-test scores by word are shown in Table 4.4 below.

Table 4.4 – Pre-test word scores ordered by percentage of pupils defining each word correctly

Word	Correct Answers		Word	Correct Answers		Word	Correct Answers	
	/82	%		/82	%		/82	%
X-rays	79	96.3%	Geocentric	49	59.8%	Transverse waves	24	29.3%
Radioactive	76	92.7%	Infrared	48	58.5%	Induced current	23	28.0%
Thermal	75	91.5%	National Grid	47	57.3%	Primary coil	23	28.0%
Sound wave	73	89.0%	Hydroelectricity	46	56.1%	Steady State theory	23	28.0%
Sound waves	70	85.4%	Acid rain	44	53.7%	Beta (β) particles	22	26.8%
Thermal imaging	69	84.1%	Ultraviolet	40	48.8%	Pitch	15	18.3%
Light waves	68	82.9%	System	39	47.6%	Neutron star	14	17.1%
Renewable energy resources	68	82.9%	Electromagnetic radiation	36	43.9%	P waves	14	17.1%
Tsunami	67	81.7%	Non-renewable resources	36	43.9%	Focus	13	15.9%
Electrical current	59	72.0%	Elastic potential	33	40.2%	Slip rings	11	13.4%
Tidal power	58	70.7%	Protostar	33	40.2%	Step-up transformer	10	12.2%
Supernova	57	69.5%	White dwarf	32	39.0%	Infrasound	9	11.0%
Microwaves	56	68.3%	Chemical potential	31	37.8%	Mean score	41.5	50.6%
Tectonic plates	56	68.3%	Ionising radiation	31	37.8%			
Kinetic	55	67.1%	Converging Lens	30	36.6%			
Direct current (DC)	54	65.9%	Real image	28	34.1%			
Alternating current (AC)	53	64.6%	Seismic waves	27	32.9%			
Wavelength	50	61.0%	Sonar	27	32.9%			
Electromagnetic spectrum	49	59.8%	Main sequence star	25	30.5%			

Baddeley, Thompson and Buchanan (1975, p.575) suggested an inverse relationship between memory span and word length. Wellington (2000b, p.142-3) described the use of polysyllabic word calculations in various reading tests, suggesting a link to difficulty for pupils in reading a text. However, there was no significant correlation between either word length or the number of syllables in a word and the pupils' ability to define the word using the post-test. This highlighted the complexity of scientific language and the dangers of using general measurements of word complexity to assess language difficulty. This is not to say that in a text the number of longer words may not be linked to reading difficulties but that such relationships may be compounded by other factors.

The decreasing order showed no apparent link between particular topics or word groups and the score. For example, pupils were very familiar with "X-rays" (electromagnetic radiation), "radioactive" (ionising radiation), "thermal" (energy), "sound wave" and "waves" (waves). They were not scoring well on "infrasound" (sound), "step-up transformer" (electricity), "slip rings" (electricity), "focus" (waves), "P waves" (waves) and "neutron star" (the Universe). This suggested that pupils knew some vocabulary in various topics but were not necessary familiar with the full range of vocabulary in each topic. This would indicate the building of additional vocabulary at each stage of learning. There did however appear to be a pattern linked to how common words are in normal language. "Infrasound", "transformers", "slip-rings", "earthquake foci", "P waves" and "neutron stars" would have been unlikely feature in day-to-day life. However, "x-rays", "radioactivity", "thermal", "sound wave", "sound waves" and "thermal imaging" would have been more common in various media and were likely to be familiar in their daily lives. This potentially

linked again to the comments by Ross, Lakin and Callaghan (2000) who said that more familiar words might have been less confusing for pupils when they encountered multiple meaning for those words. This was speculative at this stage but could be a line of future research.

The statistics also generated a list of words which required more work for pupils to acquire irrespective of reason and this could be useful for planning future teaching and interventions. This was a useful outcome irrespective of the findings and in a longer study the impact of this information could have been studied further, especially given the developing link to revision already discussed.

4.4 Familiarity of vocabulary

The *Key Stage 3 Framework for teaching science: Years 7, 8 and 9* (DfE, 2002, pp.73-80) provided a vocabulary list by year. Correlation of that list with the keywords in this study showed only four words that are in the Key Stage 3 list (with waves and light used as well in a slightly different form). Those words varied in score and, in combination with the small number of them, meant that there was little to draw out from this line of inquiry. However, it did add weight to the purpose of this study, showing that so little of the advanced Key Stage 4 vocabulary had foundations in Key Stage 3. As Academy Schools and Free Schools are not bound by the National Curriculum this concept would be complex to investigate in terms of the National Curriculum and would need an understanding of the prior learning (in terms of vocabulary) in individual schools. In the school where this research took place change to the school schemes of work at this level had been significant since the

end of SATs testing at this level in 2010. Pupils had followed a scheme of work designed to introduce them to the GCSE by studying the earlier topics at a more basic level (topics on space and the EM spectrum) in Year 9. The Key Stage 3 course followed in Years 7 and 8 by these pupils was designed “in-house” but was loosely based on the Exploring Science text books (Levesley *et al.*, 2002a; Levesley *et al.*, 2002b). Using these as a guide it was possible to construct an approximation of the year in which pupils were introduced to the keywords in the sense of their conceptual purpose in the GCSE course. A small number of these words were also common to other subjects or in regular media use. These words were grouped to give a better understanding of familiarity in the host school but it would be near impossible to be sure of the experiences of different classes and pupils, with factors such as not all topics being covered, varying approaches by different teachers and pupil absence. The groupings are shown in Table 4.5 below.

Table 4.5 – Keywords grouped by possible previous pupil encounters

Word group and description	Words in this group
May be used regularly outside of science lessons (such as in the media).	X-rays, tsunamis, light waves, sound wave, sound waves, radioactive
Used in Year 7/8 onwards and may be used significantly in other subjects, particularly geography.	Renewable energy resources, hydroelectricity, tidal power, non-renewable resources, acid rain, chemical potential, pitch, tectonic plates, electrical current
Used in Year 7/8 onwards (and not likely to be met elsewhere).	Thermal, kinetic, chemical potential, thermal imaging
GCSE words that may have been encountered in Year 9 (and not likely to be met elsewhere in the same context).	Geocentric, microwaves, infrared, ultraviolet, electromagnetic radiation, electromagnetic spectrum, protostar, main sequence star, white dwarf, supernova, neutron star, steady state theory, National Grid, system
Words not likely to have been previously encountered.	Real image, seismic waves, wavelength, ionising radiation, radioactive, beta particles, infrasound, sonar, focus, P waves, induced current, direct current, slip rings, alternating current, step-up transformer, primary coil, elastic potential, converging lens

Table 4.5 was constructed as a hierarchy in that the top category was likely to be most familiar, either in terms of length of time or range of settings, through to the words at the bottom which might have been completely unfamiliar to some pupils. A flaw in this consideration again linked back to Ross, Lakin and Callaghan (2000) in that some of the words or elements of them might have been familiar to the pupils with alternate meanings. This would have been very individual and again furthered the need for an in-depth qualitative study into pupils understanding of the individual words. Whilst a level of variance was likely with this hierarchy it was used to construct an arbitrary numerical scale of familiarity in order to analyse the relationship between familiarity and word score. This scale is shown in Table 4.6 below.

Table 4.6 – An arbitrary scale for familiarity of science keywords in GCSE physics

<u>Scale number</u>	<u>Rationale</u>	<u>Words in this Group</u>
1	May be used regularly outside of science lessons (such as in the media).	X-rays, tsunamis, light waves, sound wave, sound waves, radioactive
2	Used in Year 7/8 and may be used significantly in other subjects, particularly geography.	Renewable energy resources, hydroelectricity, tidal power, non-renewable resources, acid rain, chemical potential, pitch, tectonic plates, electrical current
3	Used in Year 7/8 and not likely to be met elsewhere.	Thermal, kinetic, chemical potential, thermal imaging
4	GCSE words that may have been encountered in Year 9 or outside of science lessons.	Geocentric, microwaves, infrared, ultraviolet, electromagnetic radiation, electromagnetic spectrum, protostar, main sequence star, white dwarf, supernova, neutron star, steady state theory, National Grid, system
5	Not likely to have been previously encountered.	Real image, seismic waves, wavelength, ionising radiation, radioactive, beta particles, infrasound, sonar, focus, P waves, induced current, direct current, slip rings, alternating current, step-up transformer, primary coil, elastic potential, converging lens

The table applied a simple numerical scale to the categories, with a “1” indicating a high likely familiarity and “5” indicating the lowest likely level of familiarity. The scale was then subjected to further statistical analysis. The familiarity scale was compared

with word score (post-test). Spearman rank correlation coefficient was used for this nonparametric data correlation of two variables (Hinton, 1995, p.271-2). A correlation of -0.7 was found. Given the high value of N (50) this result has a high level of significance (well above $p=0.01$). This was an inverse relationship meaning that the more familiar a word was (a low score), the higher the word test score (post-intervention) that was achieved. This was a simplistic analysis and very specific to this setting as well as using a scale open to further refinement. However, it did indicate a possible correlation between familiarity and word score worthy of more rigorous investigation in a subsequent study. However, the high change in scores for some of the words in category five in Table 4.6 suggested that acquiring a final high score is nonetheless not necessarily linked to prior learning or familiarity for the post-test score.

These findings were in keeping with comments from the staff focus group which included the thought that “doing it [physics 1 preparation] at the end of Year 9 has definitely helped this year...” Confidence was also mentioned in relation to this. Further support was provided by concerns about prior learning that included “...as a new teacher to the school I found it harder because I didn’t know their...prior knowledge, what they’ve been taught and how they’ve been taught.” This was confirmed by another teacher who said “There are some black holes in terms of Year 10 and Year 11 in terms of what they’d done at Key Stage 3...I was expecting [with regard to waves] to just recap...but it was back to basics.” Again, the research by Funk (2012) and the ideas of Bloom (1956) supported this desire for progression in the sciences. This all contributed to the need to understand progression better in order to explain the quantitative results fully.

4.5 Pupils with English as an additional language (EAL)

There were three further points of consideration that became apparent during the use of the pre-tests and post-tests that were worthy of note and have not been covered by the previous sections. The first covered those pupils who speak English as an additional language. August *et al.* (2009, p.358) had specifically focussed on these learners and had categorised the data separately for them. They had found that their intervention had similar impact on English learners as it did on those proficient with English, in terms of effect size. This study did not cover a sufficient quantity of EAL pupils to draw statistical conclusions on the benefits for EAL pupils and whilst the development of language would logically be beneficial for EAL pupils alongside non-EAL pupils there was a possible indication that the logic may not be sound with one EAL pupil not improving at all between pre-test and post-test. This suggested that further understanding of the requirements of EAL pupils is required for an intervention to be effective for all. This also linked to the need to further consider the impact of prior learning on acquisition of vocabulary. Lara-Alecio *et al.*'s (2012, p.1003) results suggested that science reading can be significantly improved with an intervention for such pupils, adding weight to the need to understand this issue further. General conclusions would not have been reached in this study, even with a method design to do this, due to the very small number of EAL learners in the school. Further refined interventions for science vocabulary would need to be studied in settings with more EAL pupils in order to start evaluate relative impact if there was a desire to extrapolate results for a wider use of the intervention in a variety of school settings with varying student background.

4.6 Effects of the multiple-choice test

The multiple choice test was also subjected to further analysis. This was to ensure that the results were due to pupil ability to define a word and not unduly due to the options provided by the multiple choices available. It could have been that some words had two or more options that were confusing for the pupils to differentiate between whereas other options could have made it easy to deduce the correct definition. In order to do this the options selected was tallied in a table (for both the pre-test and post-test) in order to identify any anomalies, such as many pupils selecting the same incorrect answer. This data is shown below in Table 4.7

Table 4.7 – Multiple-choice answers selected for each keyword

Word	Pre-test								Post-test							
	1	2	3	4	5	M	?		1	2	3	4	5	M	?	
Geocentric	9	49	0	8	15	0	1		6	70	0	1	4	0	1	
Light waves	6	1	68	0	6	0	1		2	0	76	1	2	0	1	
Microwaves	56	4	13	1	8	0	0		69	3	8	0	2	0	0	
Real image	27	0	28	3	21	1	2		4	0	71	1	6	0	0	
Transverse waves	30	3	1	24	23	0	1		29	0	7	41	3	0	2	
Sound waves	0	70	2	4	6	0	0		0	74	2	4	2	0	0	
Seismic waves	27	8	6	3	34	0	4		78	0	1	0	2	0	1	
Wavelength	22	50	3	1	5	1	0		12	67	0	2	1	0	0	
Infrared	13	2	4	48	14	0	1		11	2	5	61	3	0	0	
Ultraviolet	8	2	40	7	24	0	1		10	6	55	4	5	0	2	
Electromagnetic radiation	1	9	12	36	21	0	3		4	7	7	57	6	0	1	
X-rays	2	0	0	79	1	0	0		2	0	0	78	2	0	0	
Electromagnetic spectrum	13	3	0	49	14	0	3		13	2	0	61	5	0	1	
Thermal imaging	4	69	0	1	6	0	2		1	76	0	0	2	0	3	
Ionising radiation	0	3	31	11	36	0	1		0	0	64	9	7	0	2	
Radioactive	0	0	1	76	5	0	0		0	0	0	79	3	0	0	
Beta (β) particles	1	22	3	16	36	0	4		9	44	2	13	11	0	3	
Protostar	13	5	4	33	24	0	3		19	4	1	48	8	0	2	
Main sequence star	12	25	6	8	29	0	2		13	53	3	2	8	0	3	
White dwarf	32	0	18	16	14	0	2		68	0	6	4	2	0	2	
Supernova	0	57	3	7	13	0	2		0	72	2	3	4	0	1	
Neutron star	20	10	14	6	28	0	4		16	11	39	2	9	0	5	
Steady State theory	7	23	17	1	31	0	3		6	54	10	1	7	0	4	
Pitch	10	15	35	1	17	1	3		24	20	31	0	3	1	3	
Sound wave	73	0	1	1	6	1	0		71	3	0	4	3	0	1	
Infrasound	29	9	6	3	31	1	3		17	50	6	0	6	0	3	
Sonar	13	1	5	27	32	2	2		10	0	4	54	10	0	4	
Focus	0	16	14	13	35	0	4		2	30	14	23	10	0	3	
P waves	7	7	6	14	44	1	3		8	11	5	45	6	0	7	
Tectonic plates	4	56	1	1	17	1	2		2	67	0	0	8	0	5	
Tsunami	4	67	0	0	9	0	2		6	72	0	0	0	2	2	
Electrical current	59	4	6	2	8	0	3		68	2	3	0	7	0	2	
Renewable energy resources	3	68	0	0	8	1	2		6	70	0	1	4	0	1	
Hydroelectricity	3	7	5	46	18	0	3		15	7	4	41	11	0	4	
Tidal power	12	58	0	1	9	0	2		13	57	0	2	8	0	2	
Non-renewable resources	31	0	1	36	7	5	2		45	0	0	31	1	4	1	
Acid rain	44	3	0	27	6	0	2		44	6	0	27	3	0	2	
Induced current	4	23	4	8	40	0	3		3	46	2	12	16	0	3	
Direct current (DC)	54	2	0	8	16	0	2		69	1	3	4	4	0	1	
Slip rings	8	19	0	11	41	0	3		2	10	0	48	15	0	7	
Alternating current (AC)	53	6	1	0	20	0	2		67	3	0	0	8	0	4	
National Grid	14	47	1	0	18	1	1		2	66	6	0	6	0	2	
Step-up transformer	2	10	5	8	54	0	3		7	37	4	14	16	0	4	
Primary coil	9	4	3	23	37	2	4		8	8	1	49	12	1	3	
Thermal	1	2	0	75	3	1	0		0	1	0	72	5	0	4	
Kinetic	2	3	55	6	13	0	2		2	8	65	0	4	0	3	
Chemical potential	1	14	6	31	27	0	3		3	6	6	53	11	0	3	
Elastic potential	12	33	2	1	32	0	2		11	51	4	0	11	1	4	
System	10	0	0	39	29	0	4		3	0	2	57	15	0	5	
Converging Lens	19	30	2	2	25	1	3		33	36	3	0	9	0	1	

The table showed the correct answers highlighted in green. Analysing this table further showed that on some questions pupils had split their answers in rough proportion across some of the options, whereas in a small number of cases a large number had opted for a particular incorrect answer. For example, for “real image” 27 pupils had selected “an image which we can’t actually see” and for “transverse waves” they had chosen “a wave where the energy travels in the direction of displacement.” In these instances the pupils seemed to have some idea of the answer but were not familiar enough with the definition to find it from phrases which were similar but had an important error within them. This was very well exemplified by 31 pupils selecting “fossil fuels (coal, oil and gas)” as a definition for “non-renewable resources”. This led to the conclusion that pupils had some familiarity with the words but had not fully acquired the meaning of them. At this stage of analysis it was apparent that some of the words with very high levels of success attached to them had potentially less confusing alternative options. For example, “thermal imaging” was successfully defined as “making an image based on the heat given off by something” by 69 of the pupils and the other options of “a camera negative”, “using a hot knife to make an image” and “a type of insulating blanket” were perhaps too unlikely to be correct as they were not ideas discussed during the course and may have been obviously erroneous. The checklist of Cantor (1987) needed to be reapplied in light of these first results from the test. It was clear that in some instances pupils had used some level of deduction through chunking of words which would further explain errors where options were hard to differentiate in this way. For example, 14 chose “a large map that is split up into squares to show where everything is.” This seemed a logical guess when confronted with the words “national” and “grid” and no real idea of what the National Grid really was.

Where the possible answers were simply different but not clearly related to the word being defined the answers were often more equally spread. “Focus” was not defined by any pupils as “the nearest town to an earthquake” but almost equally was defined by pupils as “the place on the surface above the point where an earthquake starts”, “the place where P and S waves are furthest apart” and “the point where a seismic wave begins”. Svensson et al.’s (2009, p.205) concerns about pupils’ true understanding of words were potentially apparent here even after the intervention.

The design of the test had taken account of the need to test for misconceptions and common confusions in definitions, in keeping with the thoughts of Svensson et al. (2009, p.205). As the test was used in a GCSE course it did seem inappropriate to test this way but it did mean that care was required when considering the results quantitatively since some words were closely related to other words with different meanings (such as converging and diverging lenses) whereas others (like chemical potential) had a definition which was less likely to be confused with another term in a topic. These observations pointed to a different way of categorising data for teaching. Those words with very precise and easily confused definitions required more attention whilst pupils could learn to use their vocabulary skills to deduce other more simple definitions. However, Campbell’s (2012) advice concerning the review of those answers that were commonly incorrectly answered by pupils who otherwise scored high results needed to be applied before reuse of the test. “Non-renewable resources” seem to be in particular need of this review.

In returning to the literature at this stage it also became apparent that the design of the test could have been further utilised to encourage pupil retrieval processes and thus act not only as a test but a learning tool (perhaps countering the concerns of President Obama (2009)). Little *et al.* (2012) discussed how the design of the test could be considered in order to best achieve this outcome by using “competitive incorrect alternatives.” This would be a useful development point for any further use of this multiple-choice test.

4.7 Teacher effect

The analysis of the primary research in this study did not attempt to take account of the effects of the teacher nor the presumed ability of the pupils away from word knowledge (as has been discussed in the previous chapter). This was partly due to validity of statistics – the study was not large enough for an analysis where the groups involved contained such a small number of pupils. However, the aforementioned ethical concerns were also important. August *et al.*'s (2009, p.271) findings included a variance between groups which was potentially linked to variation in teaching styles and certainly did not appear to be ability linked. Lara Alecio (2012) opted not to take account of teacher affect. The research could have involved a larger more experimental test in order to consider this variable more precisely. Developed tools could also have encouraged more consistency in approach though this was originally discounted for a variety of reasons including the realities of accomplishing this in future studies or uses of the intervention tool. Given that teachers will inevitably utilise their own styles the former option would possibly have been more useful in understanding the effect of the teacher rather than immediately

seeking to moderate this variable before understanding whether this was necessary. It would also be a more ethical approach in its consideration of the staff acting as research participants. Indeed Lara-Alecio *et al.* (2012, p.995-6) used a teacher development approach – although one which was heavily managed and perhaps still not indicative of intervention across a wider sphere of schools. Regard was given to OFSTED's (2013, p.15) warning against over bureaucratic lesson planning and the lack of focus on pupil learning within lessons.

4.8 Concluding points regarding the first research question

Pupils did improve their score on a word test over the course of the module when the keyword sheets were accompanied by the other teaching methods. T-test analysis of pupil scores disproved the null hypothesis that there is no link between teaching pedagogy of physics and pupil scores ($t = 90.8$, $df = 51$, $p > 0.05$). This in turn statistically supports the idea that the pedagogy is responsible for an increase in scores, though it was not clear how much of this improvement was due to the intervention as opposed to the normal teaching of the module or perhaps changes in teaching in response to the intervention when considering the statistics alone. The focus groups indicated that the word sheets were an element of this improvement for some of the pupils – possibly those with the supporting levels of confidence and motivation. However, this was not the case for all pupils, with some not finding the word sheets helpful.

In terms of the word scores, statistical analysis showed a 19.6% increase in the mean number of correctly defined words using a word test between pre and post-test

with standard deviation falling 6.3%. However there was significant variation between pupils and individual words. There can be some inference made from the focus groups that the variance could be linked in part to the rigour with which the word sheets were used or other teaching-related factors though this cannot be corroborated with the gathered evidence. The data tended to support Lara-Alecio *et al.*'s (2012, p.1004-5) findings that “underscore the importance of implementing direct and explicit vocabulary instruction...” for the physics topics in question.

This research demonstrated the difficulties that pupils can have in acquiring vocabulary in science and that a change in pedagogy, including a set of keyword focus sheets, can be of use in tackling this issue. The results supported Mason, Mason and Quayle (1992), Funk (2012) and Bloom (1956) in that there was a clear need for language development and such development appeared to be linked to progress. However, it was hard to envisage a workbook system alone being able to replicate the increase in GCSE results in Mason, Mason and Quayle's study.

The findings did not allow for analysis of impact on examination results. Further understanding of the impacts of wider vocabulary and grammar skills were needed with relation to the GCSE examinations if such results were to be linked largely to language skills. However, OFSTED's (2014, p.3-4) expectations were that literacy skills became integral to modern science teaching and the intervention had provided better understanding of how this could be accomplished, albeit with further developmental work required.

Furthermore the selected definitions of the key words remain open to further challenge given that many science words can have multiple meanings, as discussed through the study by Ross, Lakin and Callaghan (2000). This furthers the importance of recognising that this study is in the context of particular definitions used at the GCSE level of study and within the current topics and specification given by the EdExcel examination board. Some words still remain open to other definitions within the courses the pupils are studying. This is a recognised issue in science that requires further consideration, not least for pupils with weaker literacy skills (such as EAL pupils) but does not detract from the findings of the survey element of this research (furthering again the benefits of the mixed-methods approach) nor the overall indications of the difficulties pupils have with the language and the variation of this difficulty between individual pupils.

The quasi-experimental nature of this research also meant that it did not fully address Lara-Alecio *et al.*'s (2012, 988) concerns over a "lack of randomized trial studies." Similarly this study failed to address Wallace's (2012, p.295) call for further investigation of whether language is "influenced by social context." Expansion of the quasi-experiment to incorporate socio-economic backgrounds of the pupils could have provided some insight into this area, although the numbers of pupils were relatively low in terms of the likelihood of gleaning statistically valid data with sufficiently high levels of probability. Nevertheless a focus on vocabulary was seen to be both necessary and useful. Furthermore this overall approach had led to improvements in pupil understanding which was in keeping with findings from Texas, USA and work conducted in New Zealand, suggesting that it could be reapplied to the wider educational community.

Chapter 5 - Analysis of Research Questions 2 and 3

5.1 Introduction

The first research question was answered using the quantitative data predominantly whereas questions 2 and 3 were more suited to use of the focus groups data, supported by quantitative data where it was useful for gaining further support or insight.

5.2 What worthwhile benefits do pupils perceive from the change in pedagogy?

Pupils' responses to the change in pedagogy, focussed predominantly on the keyword sheets, were based on three main elements: their perceptions of the keyword sheets, their perceptions of the use of those sheets in the lessons and their response to the keyword test. Additional comments that did not relate directly to the intervention were analysed separately in Chapter 6.

5.2.1 Pupil perceptions of the keyword sheets

The predominant change to teaching of the physics module was the use of the keyword sheets. As such pupil feedback directly related to the keyword sheets was considered first of all in attempting to answer this second research question. Pupils did seem to agree that they were useful, with their comments well reflected by examples such as "...the sheets do help...", "...they're always useful..." and "...they're good for...a quick recap if you've got a reasonable understanding of what

the word would mean.” However the pupil comments did make clear that there was variance in the use of the keyword sheets and that not all pupils had completed them as thoroughly as others. The quantitative data supported this with the aforementioned variance in pupil scores in the post-test and the variance in improvement overall. Another pupil said “...they [the keywords] are a good help to have in the lesson.” This demonstrated a shared opinion by several pupils that they found the use of keywords beneficial for the learning across the lesson and did not regard it as an isolated activity within the keyword sheets. This correlated with the aforementioned statistical increase in pupil scores overall.

In terms of the structure of the keyword sheet, negative comments could be summed up entirely by the comment “It seemed a bit pointless the paragraph one did to be honest.” Pupils almost universally seemed to feel the written output section was neither useful nor enjoyable. This showed the need to develop written activities to ensure that they are enjoyable and that students can see the value of them, rather than dismissing them given the established value of output in the literature. The pupils indicated that they liked the top section (the listening/reading input) saying “I like the top section,” and “I like finding out from the paragraph...” It should also be noted that a suggestion was made that “it would be a good idea to have the key points already printed off so you haven’t got to copy them all really fast.” This supported the basic idea of the reading input section and tied into the discussion points about teaching style previously made. Taken apart from other claims, this was in keeping with Nation’s (2007, p.2) input/output approach. It would seem that the vocabulary did need to be provided early in the lesson – perhaps in a variety of ways depending on the lesson – before being used within lesson tasks to develop it

further. The issues here highlighted that word development is required in keeping with the less popular sections of the keyword sheet and whilst they have not been successful in that format there is a need to develop pupil word use beyond the simple definitions so that they can become more familiar with the definition in context. This again supported those ideas of Bloom (1956) and Funk (2012) in terms of a planned development of learning.

Revision also became key, highlighted by the pupil who said "...I can't learn from [keyword] sheets because I know...I won't go back and read them." This was particularly interesting as it highlighted pupil awareness of the need to recap work though this was not an intentional part of the use of the word sheets. It also showed a difficulty pupils can have in revising or at least with motivation towards revision. The use of keyword sheets as a tool for later revision could be further considered. Further investigation could identify whether pedagogy can be effectively altered to encourage pupils to see word tasks as a learning activity and not a revision exercise should this be deemed useful in countering this specific objection to the sheets. As already discussed, the quantitative data supported further research into this area, though it was not possible to precisely compare the fluctuations in results to the level of revision that pupils had completed independently. This would have required a measurement for revision which would need to be reliable – such as specific revision tasks checked by the teacher. This would be much more involved for staff but would be a worthwhile area of study. Further consideration of revision in general is made in Chapter 6 where comments made by staff and pupils that did not directly related to the intervention were considered in terms of their wider application.

5.2.2 Pupil perceptions of the structure and use of the keyword sheets

In terms of the way in which the sheets were used, a pupil said that they "...find working with someone helps me because I learn off them as well." This supported the concept of using paired work and Wellington and Osborne (2001, p.59-60) corroborated the idea of coaching and practise in reading. Mason, Mason and Quayle (1992, p.347) fully concurred with the idea of paired work but alongside independent work too. Part of the desire to include paired work was to ensure that input/output activities were balanced although the pupil concept was that they could support each other which could be considered a similar outcome. The exact nature of how pupils supported each other would need further investigation, perhaps through more detailed and measured observation of the activities taking place. However another pupil said "I'd rather just like have one sheet for one lesson...otherwise there's like so many and I'm like so confused." The sheets were not suitable for simply bombarding pupils with as a revision or learning tool independent of a well-structured lesson. This finding does seem to conflict with the approach endorsed by Mason, Mason and Quayle (1992) given their use of workbooks for developing language skills, as opposed to segregated activities. Pupils have at times been noted to dislike excess use of worksheets for general teaching and it was clear that caution would need to be used when it came to designing interventions around such an approach. This suggested that looking at approaches that do not rely on worksheets in large numbers in future.

A more integrated use of keywords in lessons without the worksheet format was linked to the pupils' desire for active lessons, with one pupil saying they liked it when

lessons involved “Getting people up...moving around...[as] everyone stays interested then.” Another pupil said “I think the interactive side is always a lot more helpful,” These comments linked to mentions of specific modelling activities such as refraction and electricity (electron flow) as opposed to necessarily meaning practical work. Research could use this point to consider active modelling as a non-verbal tool in developing keywords. A pupil related to more general learning in saying that “I think if I’ve seen pictures or videos or something that relates...it would probably help.” This furthered the idea of learning words using a wider array of methods. Smail (1984, p.88) discussed how curriculum could be modified specifically for girls using a change in approach, such as linking physics to the human body. Further development of the teaching to meet the recommendations of Smail and other researchers of science curriculum needed further consideration for future development of the intervention.

Funk (2012, p.310) was validated by these findings as he spoke of the importance of considering “state-of-the-art approaches” with “impact on curriculum planning, distribution of learning activities, and progressional planning...” Linked to this Knipper and Duggans (2006) suggested short activities could be used as a tool in developing a better variety of writing to learn activities, furthering this idea but still suggesting that short intervention activities for keywords could have a place within lessons as per the intent of the keyword sheets.

5.2.3 Pupil perceptions of testing associated with the intervention

In considering the use of testing success there were general acknowledgements by some pupils of the importance of seeing their progress in science, such as “I think my levels have gone up pretty good.” This comment referred to science in general and not specifically the keywords or this unit but it reflected that some pupils do enjoy and respond to tracking their progress and appropriate levels of testing are a key part of the course. However, the point made by Obama (2009) that demanded further reflection on the testing was prudent at this stage. Testing was used as an indicator of progress and seemed to motivate pupils but it could not be regarded as meaning an improvement in outcomes in science without further evidence of this. The other elements of questioning in the focus group helped to ensure the wider impact in this study and reinforced the importance of the mixed methods approach.

However, as was discussed with regard to the quantitative data, the findings of Little *et al.* (2012) could have been used to develop the test so that it also encouraged recall. With careful design the concerns of Obama (2009) could have been countered and the activity would have had multiple purposes. A measured test would also help to meet Braund’s (2008) call for teachers to be aware of pupils learning when they receive a class from a previous teacher.

5.2.4 Pupil motivation towards keyword sheets

The importance of ensuring pupils were motivated was reflected in their comments about topics and motivation in science. There was broad agreement between

several pupils that interest (and thus motivation) were contributing factors with comments including “Like when you learn about the Earth and stars...I think it was because I was interested in learning about it that I took more in.” Another pupil said she liked topics that related to Geography “...because it relates to the plates and stuff of the Earth.” Despite the complexity of the concepts in these topics it is clear that where they spark interest the pupils feel their learning is stronger. Despite pupils feeling that this link is to do with relevance to their lives, the fact that they dislike electricity as a topic suggests further exploration of this idea is needed, though teachers would be very familiar with the fact that space is an almost universally enjoyed topic and electricity concepts can be more difficult. However, the results of the quasi-experiment did conflict with the pupil comments in that they did not show a pattern for vocabulary that matched topics. For example, “Direct Current” (84.1%) and National Grid (80.5%) were fairly high scoring words within electricity, a topic the pupils indicated that they did not like, whereas “Neutron star” (47.6%) and “Protostar” (58.5%) were lower in score. Whilst other words showed the opposite pattern this indicated that favourable topics did not necessarily result in stronger understanding of the vocabulary. This reinforced McAteer’s (2013) concerns about the difference between what the interviewees perceived to be true and the actual reality from observations meaning care was needed in presuming that the pupils’ comments reflected useful fact. Again the evidence pointed towards a complex set of interactions when it came to vocabulary acquirement and retention.

Despite the complexity of the interactions it was still deemed important to consider motivation further since the research did not find that it was irrelevant, simply that it was not an overriding variable when considered with other elements. One pupil

remarked that "...if you did something to make them [the lessons] more interesting people would pay more attention." As already discussed Gorard, See and Davies (2011, p.116) found that individual motivation had a medium positive effect and as such its importance should be taken into account when considering variation in pupil performance and also in considering how to improve interventions to ensure pupils are motivated to maximise their use. The issue of some pupils not being fully motivated for the tests showed the issues that could start to arise and demonstrated the complexity of motivation further.

5.2.5 Conclusions in response to the second research question

Pupils had found the increased focus on keywords useful, with many particularly finding the comprehension activity of the keyword sheet helpful. However there were some variances in utilisation of the sheets in teaching and in pupils' disposition towards them. The combined comments reflected the need to trial the word sheets in a more consistent manner, if experimental or quasi-experimental research was to be used and the results were to be more reliable. They also reflected a great variety in teaching styles and a clear feeling from pupils of activities which helped them and which did not work for them. These feelings appeared to mirror current professional advice. Further exploration of this theme could ensure that pupils genuinely benefit from the prescribed modern methods and that these are utilised to drive staff professional development and reflection. This would be in keeping with the recommendations by Lara-Alecio *et al.* (2012, p.1006) who said that the implications of their study were "...particularly related to professional development for teachers who have ELLs and economically disadvantage pupils in their classrooms..."

Concurring with this, August *et al.* (2009, p.371-2) concluded that “...professional development could be designed...to help ensure all teachers can achieve comparable gains...”

Suggestions for further development from pupils included “...maybe the complicated words maybe have the definitions in front of you so you know how to use them in the work” and “they could be like the tables you have in other lessons – the big laminated sheets in the middle of the table...like science vocab or something.”

These thoughts were expanded upon with pupils highlighting that the words also require more context to revise their meaning properly. For example, “...you’ve just kind of got the word on its own and you really need to be able to link it to other things...” This endorsed Wellington and Ireson’s (2008, p.175) recommendation for use of dictionaries or glossaries, in conjunction with highlighting new words. This could form the basis of one of a number of more varied activities to introduce (or reinforce) the learning of complex vocabulary.

An obvious complexity with regard to GCSE science was the sheer number of keywords within science. A series of topic based sheets would be required as opposed to the one sheet that some subjects have managed to produce due to the repeated language in those subjects. Further research is required to develop this tool properly and decide upon the appropriate resources. Given that keywords are given in context in revision guides and in textbooks further consideration could be given as to how best to use these resources, as has already been discussed. This would also link to consideration of long-term curriculum planning and organisation.

5.3 What worthwhile benefits do teachers of the pupils perceive from the change to pedagogy?

The teachers' comments followed a similar set of themes as the pupils: their overall feeling about the sheets, the structure and use of the sheets in lessons and the relationships to pupil motivations.

Staff talked about keywords directly more than any other topic. This was likely to be due to their awareness of this research and the principle approach behind it.

Nevertheless their comments about this topic were varied. Notably the word "confidence" was used a lot. It is clear that staff felt a lot of pupil confidence was bound up in the keywords, highlighting the idea of solid foundations that needed to be built before ideas could be developed further. It should be reaffirmed at this stage that this is in keeping with department philosophy and the use of Bloom's taxonomy was integral to this. This may have influenced staff approach subconsciously.

5.3.1 Teachers' perceptions of the use of the keyword sheets

With regard to the word sheets, staff remarked that "...once they are becoming familiar...they are becoming much happier and there's less resistance with them..."

This highlighted initial issues with changing pedagogy and also the importance of routine and familiarity. These issues were found by Adey (1997, p.7) when reporting on curriculum change in science and he felt that a period of time was required before change was accepted.

More negative comments about the word sheets included “I think I used a load in a row in lessons in a row and I must admit I did stop because they started moaning oh you know not another one of these sheets again.” This opened up the need to revisit the consideration of how directed the use of an intervention should be – instructing that the sheets were to be used one per lesson could have avoided this issue but would have lessened staff ownership of the process. One member of staff said they were of “...varying degrees of use, depending on the lesson.” This highlighted the converse need for teacher variety in approach in adapting to the needs of the pupils, which they themselves have talked about, albeit less directly. Another teacher said “I’ve just picked ones as and when I’ve felt I’ve had time...” These variations in approach were clearly not a controlling factor in quantitative results given the variation between pupils within sets with different teachers, though these comments did lend support to the approach of August *et al.* (2009) in evaluating teacher impact on the intervention, although the ethical decision not to do this still overrode this consideration.

The concept of using more varied techniques to help pupils develop the use of keywords beyond the word sheets was provided by staff and showed a strong case for allowing teachers to use their own strengths in teaching. Ideas suggested included “...trying to teach silly little ways of remembering things...” and “...*Taboo* games and bingo games...” These ideas supported the aforementioned support for variety in teaching and particularly in language development, such as by Wellington and Osborne (2001).

5.3.2 Staff perceptions of the keyword sheet layout

The reading comprehension again seemed to be the section perceived as the most useful with one member of staff saying “The top part was very, very good especially if you’re speaking to them and they’ve had to listen and they’re writing something down and then they’re checking it with reading.” Other comments included “I think they’re good. They’ve definitely helped...the top bit as a starter...the bottom bit as a plenary.” Another teacher said “I’ve mainly just used the top bits with just the keywords...” The benefits were felt according to one teacher who said that “...they’ve answered exam questions more confidently if at least they see a word and have some idea what it means...” With regard to the rest of the keyword sheets one staff member summarised that what the pupils “...don’t want to do, they don’t see the point in doing, is writing out the speech...” though noting that “...you can get them up and speaking about it.” This mirrored pupils’ feelings and suggested that spoken output can be of use but that the written output is not useful in this context – but perhaps has its place within the teachers’ lesson plans as a whole, again in line with Nation’s (2007) recommendations and the concept of planning for developmental progress in science (Funk 2012 and Bloom, 1956).

5.3.3 Teachers’ perceptions of testing in connection to the keyword sheets

Staff did not comment directly on the keyword tests but made some general points about testing and assessments. One teacher said “I think the mini-tests throughout the topic helped as well...they’re always asking ‘What does that convert to then, what grade?’” This was supported by another teacher who felt that the use of

assessment "...focussed their minds as well because they always start off poorly with the first couple then realise they can't just go in there like in Year 9 and probably get a mark at the end of it." This again suggested a need to consider the Key Stage 3 preparation for Key Stage 4. There could be a difficult balance with confidence where pupils struggled, though one staff member felt that lower grades were a good thing in that "...if they don't know it they fail...badly and get E's and F's and that really brings it home to them and it needs it, they needed that." Similarly a point was made that pupils had come asking for help to improve grades from E's and D's. This possibly linked to Goodman and Gregg's (Eds., 2010) findings that belief in actions was important. The pupils who improved clearly knew that they could improve and how to go about doing this. Again motivation was clearly a driving factor (Gorard, See and Davies, 2011; Galton, Gray and Ruddock, 2003).

5.3.4 Teachers' perceptions of pupil motivation

One member of staff felt that "there was a marked difference in the attitude between Year 10 and Year 11. Year 10 are much harder going." This could have been a reflection of experiences in Key Stage 3 but also could be due to differences in the cohort. The modern measurements of pupils levels of progress has meant that progress is relative to pupils (and a cohort) but does make it harder to reflect on overall progress since there can be an expectation for learning to follow similar patterns each year, which may not be realistic but makes planning teaching more complex. This was linked back to revision: "...they're seemingly not prepared to go away as told and read through their work that night or that weekend just to keep it fresh in their minds."

There had been some debate as to whether there is a need to structure the revision for the pupils so that they are compelled to do it, such as setting homework task sheets. This thought was shared by a colleague who said "...a lot of them don't like doing homework and if you set them revision they don't like it but...my group...liked review questions because that focussed them." This showed that carefully planned review tasks were seen positively by pupils and so could be of use in structuring revision further, rather than expecting pupils to manage their revision independently.

Staff and pupils clearly indicated that motivation affected learner progress yet Galton, Gray and Ruddock (2003, p.56) found that "...the pupils who performed best often had the poorest attitudes to science." Though Galton, Gray and Ruddock (2003, p.56) do note the high level of motivation in these pupils. This could have been an indicator of differences between the Triple set and other sets – those pupils in the Triple set were encouraged to opt for extra science if they had a strong intention of studying it at 'A' Level. The importance of motivation was discussed with these pupils during their options in Year 9. Gorard, See and Davies (2011, p.116) appeared to back this claim given that across a multitude of studies they found only weak correlation between individual attitude and attainment. They did find however that individual motivation had a medium positive effect. So perhaps whilst the well-being and positivity of pupils needed to be an addressed concern it should not be confused as necessarily having an impact on attainment which remains an ultimate objective.

5.3.5 Final conclusions in response to the third research question

The teachers had concurred with pupils, finding the first section of the word sheets had worked well. Motivation concerns were repeated within other themes discussed by staff. It seemed clear that the factors affecting motivation required further understanding. They may have been very individual and certainly could have been specific to this school so further research needed to focus on psychology or at least be focussed in terms of the population and setting.

The principal use of the word sheets was to provide a starter for the pupils. The word sheets seemed to be too lengthy, not entirely useful in their entirety and were used too frequently (which could be compounded by the time spent on them). Staff also detected the pupils' desire for a glossary and one teacher said "Some of them are very keen on having a glossary so they have made a big thing of putting information in the back of their books..." This supported the developmental idea of tying this process into a glossary for each topic in some form. This could then further link to vocabulary activities during the lesson.

Mohan and Slater's (2006, p.314) called for greater understanding of the science register in science classes. It was certainly clear that like the teacher Mohan and Slater discussed (2006, p.303), each member of staff involved in this study had taken on a role of directly teaching students scientific vocabulary directly and in doing so had improved their analysis of the science register in their classes. This model responded to Reay's (2006, p.303) concerns that teacher-training did not adequately prepare teachers for tackling social class issues in that it also extended

the training of staff members. There remains further potential to impact wider study and to inform such teacher-training, both within the school and beyond.

Chapter 6 – Further Discussion

Having considered the three research questions, it was apparent from the thematic analysis of the interviews that several themes had arisen which were not directly relevant to the research questions. However, the points were still important in considering the development of pedagogy in physics and science further in the sense of the overarching research question.

6.1 Lesson content

Pupils made several comments about lesson content. Several of these comments were critical in nature. One pupil felt that "...the whole class will get it but then the teacher will just move on so it's like harder and then I have to go back for extra..." Other pupils expressed the need for additional lessons (though these were regularly available – highlighting a lack of awareness to this). This linked to wider pedagogical considerations, such as effective use of progress checks to ensure pupils have grasped a concept before developing it further or moving on. Differentiation also then linked to this since this pupil clearly required more support in class. The use of setting by ability could also have been linked into this discussion though clearly this would not detract from the need to cater for all pupils in a group irrespective of administrative decisions. Skribe-Dimec (2013, p.194) found differentiation was not used by 90% of trainee teachers in primary science. Whilst aimed at a different education phase this suggested further understanding of the use of differentiation was needed in order to ensure that it was used effectively in science lessons. This also heightened the concerns about teacher-education that Reay (2006) had raised.

Whilst the teachers in this study were not newly-qualified recent CPD had not focussed on vocabulary and as such Reay's concerns remained relevant to more experienced staff.

The progress check element was emphasised further by a pupil who said "...when you're sat there you can sort of choose not to pay attention", "...not getting it at all", "it goes a bit too fast" and "[the teacher] skips on and you have to ask [the teacher] to go back." OFSTED (2013, p.15) warn that not checking progress in an appropriate way can damage the quality of lessons. This seemed to be supported by the pupil comments, though these comments considered in isolation of what actually occurred during lessons meant that it was unclear whether a change to lesson planning was needed to support these pupils since the pupils may not have been fully utilising resources at their disposal.

6.2 Pupil motivation

The issue of pupil motivation was discussed in relation to the keyword sheets and was clearly shown when one pupil simply admitted "I don't really pay attention as much as I did before so I don't really learn that much." However, the comments widened the consideration of the issue beyond the changes applied to pedagogy by this research and so further analysis was deemed appropriate. For example, another pupil said of attempts to differentiate a lesson by having answers available for guidance, "...everyone always just goes to the side to look at the answers and maybe it would be more helpful if say they were put so that midway through the lesson you could look at them." This emphasised further the need to understand

pupils' own roles in developing their understanding in science alongside techniques for developing teachers' pedagogy. The two could not be seen in isolation of each other. Staff did not discuss this issue. This might have been due to pupils being aware of a broader spectrum of teaching styles and lesson approaches since pupils participate in learning with several staff across a week whereas teachers mostly reflect upon their own experience, viewing other teachers' lessons far less frequently than a pupil. This highlighted a potential need for staff to increase their participation in lessons delivered by other staff in order to allow them to reflect more widely. This reflection in turn could be used to develop the pedagogical tool further, with activities selected reflecting those that most motivated and interested pupils from across the school (and beyond), again using the recommendations of Smail (1984, p.88).

One pupil indicated that they had enjoyed a trip to the local college to see radioactive element investigations demonstrated. Time and cost restricted trips during the module but visits out of school were important and should not be forgotten at this level. They might have been particularly useful to support further development of vocabulary and in particular for exploring the effect of attentiveness in different setting upon the retention of vocabulary. Wellington and Ireson (2008) supported various methods of complementing learning using science outside of the classroom, noting that it should be additional to classroom-based learning and not in place of it.

The importance of interesting and varied lessons was made clear by OFSTED (2014, p.7-8) in support of this concept. Dillon's (2011, p.5) reminder that science in the media generated plenty of interest – despite the language and concepts used – also acted as a reminder that the subject itself was not necessarily a barrier to

ensuring that this happened. Perhaps television has focussed on those areas that people are naturally curious about but many elements of science could be linked to interesting stories or perspectives. The findings of Galton, Gray and Ruddock (2003) did however question the importance of attitude in comparison to motivation simply to do well, acting as a reminder of the complexity of attempting to use motivation, aspiration and enjoyment in developing interventions. Further research could have linked findings to pupils' desires to follow particular career paths or their value of science qualifications in their pursuit of future courses and employment. This further linked to impact of these issues on those elements of the subject that were less interesting by their very nature, such as revision of the work after the lessons. Gorard, See and Davies (2011) had indicated that interventions in this area had proven to have some success so further development could have included these ideas further in a wider package of interventions.

6.3 Revision

Pupils discussed revision frequently as has been discussed with regard to using the keyword sheets for revision. This highlighted their awareness of the importance of revising and their awareness of the impending examinations. In discussing ideas more widely, one pupil said "I just use revision guides. I don't find my book helpful." This was recognised by a pupil who said "...I feel like in my [exercise] books I have to go through lots of pages to get to one little point." Pupils seemed to feel that exercise books ought to have been a place to begin revision even though these exercise books were used for the purpose born out in their name – exercises in class. They were not designed to be an all-encompassing revision aid and all pupils

had purchased revision guides for this reason. Clearly the pupils had accepted this but the comments suggested that perhaps other experiences in the past or perhaps in other subjects were leading them to a different approach to revision. Other subjects used exercise books to develop essays that could arise in examinations so where consistency of approach was not possible it was important to aid pupils in knowing how to revise for science. Future development of the intervention that involved using the “four strands” concept throughout the lesson would potentially have seen useful language-based activities being developed in exercise books and this could have proven useful for revising from later.

Pupils were also often unaware of revision tools that had been created in school for their use, such as revision flash cards – further highlighting the need to train pupils in how to revise. Pupils did select a particular example of this having happened successfully, saying “...[the teacher] gives us them [question sheets] for homework to do for tests the next day which they help a lot instead of just looking over my book. I read it and it just goes...” The pupils appreciated the revision tasks, perhaps due to a relative inability to revise effectively independently at this stage, with the sheets providing a frame to build upon. The development of this revision tool showed the importance of preparing revision to the staff and unsurprisingly revision was prominent in the teacher discussions. One teacher reported that pupils had “made significant improvements of 2 or 3 grades because they’d had very specific questions to help them revise.” The question-based revision process was supported with another teacher promoting revision sheets saying “Those packets...I’ve used the Chemistry one...” Yet another teacher said “...they did seem to enjoy those

broadsheets...” noting that asking questions in different ways was helping the pupils prepare for examinations.

The success of these revision sheets further emphasised how well developed worksheet-based tools could be successful in science. The vocabulary intervention introduced ideas rather than reviewed them. Whilst comments already discussed showed that excess use of sheets for keywords was unwelcome there was potential to consider the revision of the keywords as well as their introduction. Regular repetition of learning fell within Nation’s (2007, p.10-11) and Wellington and Osborne’s (2001, 24) recommendations and this could be further built in to a more complex intervention that revisited words when topics were revisited. This approach would potentially involve many (or all) aspects of the lesson, including practical work.

6.4 Practical work

Enjoyment of practical investigative work was clear. Pupils’ comments included “I do better in practicals” and “...like when we do practicals and stuff that helps me more.” These findings were in keeping with a common finding in science. Toplis (2011, p.534) and Abrahams and Millar (2008, p.1946) concurred with this based on several pieces of research. Whilst the very different nature of practical work might have resulted in increased enjoyment with little link to the issues in other forms of classwork there were aspects of this to consider further. As already mentioned, investigative skills are used throughout the five years at secondary school and often a long time before. Vocabulary would have been developed over time, using repetition of keywords for compiling reports and using the necessary skills. This

would be more in keeping with Bloom (1956) and Funk's (2012) recommendations which had not been met in other areas of science. An approach mirroring this with factual topics would be of interest in learning more about the role of ongoing and regular progression in science. The positive attitude to practical work from the pupils also suggested that it could be used to help develop pupil dispositions towards topics, possibly helping to make those topics more enjoyable. Wellington and Ireson (2008, p.211) suggest that pupils can however learn about science by doing investigations and suggest a mixture of approaches might be best for accomplishing this. Francis, Hutchings and Read (2004) also discussed the importance of engagement in science topics. They found that practical work was important as a part of achieving this but although it was not found to be a main focus in successful departments.

One teacher felt that "getting them to understand that when they're doing practical that they're still learning..." was a challenge, suggesting that simply linking topics to practical work would not have overcome all of the issues in ensuring pupils develop their understanding in the topic. Another teacher said "I think there are a few that seem to sit around and just assume that the practical bit is a bit of a doss time rather than you know it's there for a purpose..." This might have meant that some pupils liked practical work because they were not working with as much focus and were actually enjoying other distractions at this time. More research was required in this area in order to be sure of the best approach to utilising practical work to ensure learning is taking place but also to motivate pupils and this finding is in keeping again with the work by Toplis (2011, p.546-7).

Abrahams and Millar (2008) discussed the complexity of linking practical work to developing scientific understanding and whilst the department's use of a Bloom's taxonomy (1956) approach was applicable to considering the level of learning involved in practical lessons there was still further potential within the department and seemingly in science education more widely to develop practical work and scientific learning further. A barrier to achieving this was raised by another member of staff who felt that resources were not always well functioning saying "things take longer than you expect and with better quality equipment you could probably do a lot of practicals a whole lot faster." These issues would have to be overcome in developing more practical approaches further. Given the cost of practical equipment and the need to involve the team of science staff, including technicians, a resolution to this would need to involve extensive evaluation of equipment and its use. Though not suggested implicitly, this would be in keeping with the need to reconsider the teaching of practical work as suggested by Abrahams and Millar (2008).

6.5 Mathematics and STEM in science

Wellington (2000b, p.168) had included mathematical vocabulary as level 4 (the highest level) in his taxonomy of science vocabulary. The keyword sheets did not contain mathematical words to avoid conflict with whole school approaches but pupils did comment widely on mathematics as an issue and it may be that the keyword intervention could have been extended to help them with the vocabulary aspect. Pupils were however more aware of skill based difficulties than vocabulary-related ones (although they may have had issues with both). Of particular interest one pupil said "...them little equations things. They confuse me so much, they just

go through my head, I don't remember any of them, not one." Pupils were not required to memorise physics formulae in the EdExcel physics 1 module at this time (though this was expected to change for 2017 examinations). Pupils did also express concerns with rearranging equations and using maths for measurements and in chemistry 1 as well. These points suggested that more time should be spent on the mathematical skills of physics. One pupil said "It would be good to have a lesson on that just to get your head around it." This was echoed by other pupils who wanted more time to explore the skills involved. Pampaka *et al.* (2012, p.474) also suggested methods to address some of the issues with mathematics learning which would coincide with the pupil requests for more time dedicated to developing these skills. Such a change in approach would require more lesson time or again a review of the structure of science over five years at high school.

Pampaka *et al.* (2012, p.473) had concerns for mathematics and "teaching-to-the-test and its association with declining dispositions towards...STEM subjects at university." When this research is placed within the wider sphere of science education and STEM education it is important to consider the wider impacts as well. This would very likely have required a longitudinal study – perhaps several studies that began by further understanding the importance of vocabulary in science in a wider context and with better tools, before including the wider teaching and learning picture and then considering the combined effects on the subsequent choices made by students affected. This would likely need consideration of the impact of STEM activities, such as those suggested for girls by Stage *et al.* (1985).

6.6 Lesson scheduling

Pressures for curriculum time meant that considerations for developing mathematics further and better utilisation of practical activities were balanced against the time available in the sense of the one (or occasionally two) hour lessons and the four hours per week (or six-and-a-half for Triple science) that was available for science. Teachers commented several times about the pressures felt. One teacher said “I only see them four times over two weeks and two of those are on Friday afternoon which is not a good time...” The same teacher said “Last year I had all the GCSE groups in the mornings...a lot easier...” Another said “We seem to lose kids constantly with things...then music lessons.” These issues were relevant to the deeper understanding of variation between pupils but are an unsolvable part of the complexities of timetabling and the wider school experience. They are not followed further in this analysis since they are within an area that staff can flag up but changes would have to be considered by a senior management team in the context of impacts elsewhere. In a sense the issue here is a conflict between subjects which will always be complex and has no perfect solution.

6.7 Concluding points regarding further discussion

The themes that had arisen from the focus groups had provided a wider insight into the areas that pupils and staff felt affected pupil outcomes. Whilst the intent had not always been to discuss these points in relation to the keyword intervention it became clear that the understanding of these issues was useful for further analysis of the intervention and considering its future development. The intervention needed to

consider motivation and interest, be used effectively within the lesson as a whole, take account of revision as well as the prior learning within a topic, consider the best use of practical work for motivation, engagement and learning (of practical skills, terminology and wider scientific ideas) and potentially extend to mathematical difficulties as well. Further consideration of the use of lesson time would be imperative, with a clear structure for the five years at secondary school in keeping with the findings in the previous chapters. These issues all had the potential to have an impact if the intervention was reapplied or redeveloped for use in another setting and these insights should be utilised to help maximise its impact.

Chapter 7 – Final Conclusions and Future Directions

7.1 Introduction

The research addressed each of the three research questions in turn, utilising the findings from the quasi-experiment and the survey. The two research frames complemented each other in terms of giving a fuller picture that neither could have provided alone. This allowed for a final review of the over-arching question and final comments on those areas of analysis more suited in considering the research as a whole, particularly in responding to the literature that called for future studies to respond to areas such as effect size and cost.

7.2 Are there discernible benefits to pupils in physics when science pedagogy includes vocabulary-based activities?

The three research questions showed that science lessons had ensured that pupil vocabulary was developed for most pupils. The quasi-experiment had shown that a small number of pupils had not developed at all and this was linked largely to EAL issues and a small number of pupils not motivated to engage with the post-test, due to undefined issues that could have included their own aspirations and confidence. These pupils required an intervention that sought to address these issues. The focus group interviews supported the use of the keyword sheet intervention in helping some of the pupils in this improvement but not all pupils felt that it was helpful. There was a possible link to the way in which the sheets were used, particularly with regard to revision later. There was also a greater potential benefit

for those pupils with some prior knowledge but not as much at the top end of the range of pre-test scores.

Lara-Alecio et al. (2012, p. 1002) found that although pupils improved their reading ability they did not improve their science achievement marks. This served as an additional warning against assuming that improved vocabulary would result in improved exam scores. Therefore there were benefits to the vocabulary of most pupils but the wider benefits were not defined in this study.

7.2.1 Effect Size

Gorard, See and Davies (2011, p.121) recommended “development work on AABs [attitudes, aspirations and behaviour] and post-16 participation” and to “encourage a move towards reporting of effect sizes, and where possible, the costs of interventions”. AABs were not directly studied but were found to have impacted on the research and required further analysis in order to fully understand their impact on the pupils. This included understanding why a small number of pupils were not engaging with the process (and attempting to encourage them to do so) and the role of attitudes and aspirations linked to revision of topics and improvements over time which could have been related to pupils’ belief in their ability to improve.

The research did not provide a clear effect size due to the nature of the quasi-experiment and the relatively small population, although there was a measurable improvement in pupil scores of +9.9 words (with a +1.0 increase in standard

deviation) for 82 pupils which could be used for comparison to other studies and an indicator of potential effect size prior to an experimental study.

The survey analysis of the keyword sheets seemed to uphold the ethical decisions made in the sense that the intervention required refinement and further evidence of its value, perhaps through action-research, before a wider experimental test could have justified the overriding need for a control group and regrouping of pupils for an experimental approach. This meant that Gorard, See and Davies's (2011, p.121) encouragement towards reporting effect sizes was considered but was not met at this stage in the research but ought to remain a future goal of suitable subsequent studies.

7.2.2 Cost of Interventions

Gorard, See and Davies (2011, p.121) also suggested reporting the cost of interventions. This was easily estimated as the only material cost was that of printing the keyword sheets and tests. This was minimal and well within the day to day expenditure of a science department – though it may be that some departments would have preferred to avoid the photocopied sheets and to look for a way of providing the task from a projector or computer. Schools with tablets or other electronic devices could have easily adapted the tools to use electronically and thus eliminated materials costs. A rough estimate placed a cost of 23 word sheets plus the four pages of the test (twice over) at 31 multiplied by copying costs (around 4 pence per sheet). This brought a total of around £1.24 per pupil. Staff time for CPD and potentially a CPD expert could have been considered as without cost or valued

according to the price of hiring an expert and the cost of teacher's finite time but this would be more variable according to needs. In this study the time used was within normal meeting time and as the researcher led the department there was no expenditure for this process.

7.3 Future research directions

These combined considerations in combination with the wider findings led to a need to use the findings here alongside other research in order to develop and measure a wider ranging intervention.

7.3.1 Development of the intervention

It was clear from DfE (2013c, p.5) publications that language would continue to become more complex in science. Teachers would have to continue to explore different avenues in order to help pupils to access and progress within Key Stage 4. OFSTED (2013, p.12) were also committed to seeing this happen in lessons.

Lee and Fradd (1998, p.18) suggested – more research was needed, particularly into the cultural and linguistic variations between pupils. This could well have involved consideration of teaching longitudinally with a study perhaps over five or more years and its impact upon different groups of students.

The longitudinal approach would also lends itself to the recurring need to ensure that teaching and curriculum planning led to a levelled approach that developed

vocabulary, mathematics and scientific ideas over time (Anderson and Sosniak, 1994; Bloom, 1956 and Funk, 2012). This approach also needed to consider appropriate teaching and learning developments and ways to engage pupils (in this case girls) in lessons – both at a departmental planning level and whilst planning individual lessons (Archer *et al.*, 2013; Dillon, 2011 and Francis, Hutchings and Read, 2004). The individual lessons would utilise the recommendations of Nation (2007) in utilising balanced input and output of language but throughout lesson activities and with less use of a worksheet approach. This developmental curriculum would require development of prior knowledge of pupils' abilities and acquired knowledge which the test could aid (Galton, Gray and Ruddock, 2003). Practical work could also be considered further to ensure that pupils are motivated through it but also to ensure that the required learning links are made (Toplis, 2011 and Abrahams and Millar, 2008). The role of revision could also be considered further, perhaps through a link to parental involvement as was the case in the Lara Alecio *et al.* intervention.

These developments would require ongoing measurement (possibly through a mixed methods approach) in order to monitor their success individually or as a combined approach. Learning points from this research could be used to develop these measurements further.

7.3.2 Development of the word test

The word test was largely successful, particularly in terms of length and the data produced. However, those multiple-choice options which in hindsight had potentially

led to easier questions needed to be reconsidered in line with Cantor (1987) and Campbell (2012). Little *et al.* (2012) could also be revisited in order to fully utilise the test as a revision aid as well as a measurement tool.

7.3.3 Development of the survey interviews

The survey questions provided considerable useful data but they did fail to consider the opinions of EAL pupils whose parents spoke another language and those pupils not engaging fully with the test. Any subsequent focus groups should try to address this problem perhaps by considering wider methodology beyond interviews.

A larger gap between focus groups, as advised by McAteer (2013), would have been useful in fully analysing focus group data and then returning to the participants in order to address questions arising that had not been addressed. There would be a danger in this resulting in leading questions. An approach whereby follow-up focus groups took place could be useful in gaining an initial set of data before allowing the risk of leading questions to increase in order to refine the data gathered. This would have allowed areas such as the reasons for pupils' enjoyment of practical work and learning in that area to be considered further. Alternatively an action-research design frame would have allowed for refinement of methods as well as the intervention as the study progressed. This approach would be suitable for a longer period of research.

7.4 Summary

This study gave rise to many further questions and development points for teaching science, not only in relation to keywords in GCSE physics but to the wider need to create a considered programme of science taking into account various elements of teaching and learning and pupil psychology. In doing so it found that these points are in common with research in other areas of science teaching both in recent studies and also going back to Bloom's work in 1956. In addressing this twenty years ago, Anderson and Sosniak (1994, p.200) referred back to Bloom's (1956) own conclusions in saying that "...perhaps whoever is involved can benefit from the experiences of those who developed the original version and create a new heuristic structure that is even more successful and useful than this one has been." This seems true still today. Furthermore it would be true within similar schools in the United Kingdom – working class girls' schools – but potentially also to other school, including those internationally as the research from Texas, USA in particular had demonstrated. The research also opened up the need to further evaluate teacher-training and CPD to include the importance of vocabulary development and the needs of pupils from varying socio-economic backgrounds. Whilst this small scale research benefitted an entire department there remained a further need to ensure that research in science education continued to build using such experiences alongside larger studies and evaluative work in order to lead to ongoing progress in teaching and learning in science (and other) subjects.

List of References

Abrahams, I. and 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), 1946-1969.

Adey, P.S. (1997) *Factors Influencing Uptake of Large Scale Curriculum Innovation*
In: Annual Meeting of the American Educational Research Association, Chicago, Illinois, USA, 24th to 28th March.

Anderson, L.W. and Sosniak, L.A. (Eds.) (1994) *Bloom's Taxonomy: A Forty-Year Perspective*, Chicago, NSSE.

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, and Wong, B. (2013) 'Not girly, not sexy, not glamorous': primary school girls' and parents' constructions of science aspirations, *Pedagogy, Culture & Society*, 21(1), 171-194.

August, D., Branum-Martin, L., Cardenas-Hagan, E. and Francis, D.J. (2009) The Impact of an Instructional Intervention on the Science and Language Learning of Middle Grade English Language Learners, *Journal of Research on Educational Effectiveness*, 2(4), 345-376.

Baddeley, A.D., Thompson, M and Buchanan, N. (1975) Word Length and the Structure of Short-Term Memory, *Journal of Verbal Learning and Verbal Behavior*, 14, 575-589.

BERA (2011) *Ethical Guidelines for Educational Research*, London, BERA.

Bloom, B.S. (Ed.) (1956) *Taxonomy of Educational Objectives: Book 1 Cognitive Domain*, London, Longman.

Borg, W.R. and Gall, M.D. (1989) *Educational Research: An Introduction* (5th Ed.), New York, Longman.

Braund, M. (2008) *Starting Science Again*, Sage, London.

Brown, P. (1987) *Schooling Ordinary Kids*, Tavistock, London.

Burns, J. (2013) *Pupils behind at seven fail to catch up, warns charity*, BBC News online [last updated 8th October 2013], available at:

<http://www.bbc.co.uk/news/education-24434515> [Last updated 8th October 2013]

[Last accessed 29th October 2013].

Campbell, D.E. (2011) How to write good multiple-choice questions, *Journal of paediatrics and child health*, 47(6), 322-325.

Campbell, D.T. and Stanley, J. (1963) *Experimental and Quasi-experimental Designs for Research on Teaching*, Boston, Houghton Mifflin, available at

http://moodle.technion.ac.il/pluginfile.php/219643/mod_resource/content/0/Campbell_and_Stanley_1963.pdf [Last update unknown] [Last accessed 30th October 2013].

Cantor, J. (1987) Developing Multiple-choice Test Items, *Training and Development Journal*, 41(5), 85-88.

Carlos, M.S., August, D., McLaughlin, B., Snow, C.E., Dressler, C., Lippman, D.N., Lively, T.J. and White, C.E. (2004) Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms, *Reading Research Quarterly*, 39(2), 88-215.

Cohen, L. and Manion, L. (1994) *Research Methods in Education (4th Ed.)*, London, Routledge.

Cohen, L., Manion, L. and Morrison, K. (2007) *Research Methods in Education*, Abingdon, Routledge.

Corte, E., Verschaffel, L. and Van De Ven, A. (2001) Improving Text Comprehension Strategies in Upper Primary School Children: A Design Experiment, *British Journal of Educational Psychology*, 71 (2001), 531-559.

Creative Research Systems (2012) *Sample Size Calculator*, Creative Research Systems, available at <http://www.surveysystem.com/sscalc.htm#one> [Last update unknown] [Last accessed 31st October 2013].

De La Rouviere, N. (2013), *The Four Strands of Language Learning*. *Confuse Laowai*, [blog] 20th June 2012, available at: <http://confusedlaowai.com/2012/06/four-strands-language-learning/> [Last update unknown] [Last accessed: 1 Oct 2013].

DfE (2002) *Key Stage 3 National Strategy Framework for Science: Years 7, 8 and 9*, available at: http://dera.ioe.ac.uk/5308/1/sc_fwkd1_10.pdf [Last update unknown] [Last accessed 27th May 2014].

DfE (2013a), Department for Education Services and Guidance, accessed online at: <http://www.education.gov.uk/> [Last update unknown] [Last accessed on 26th September 2013].

DfE (2013b), *Reforming the accounting system for secondary schools: Government response to the February to May 2013 consultation on secondary school accountability*, available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/249893/Consultation_response_Secondary_School_Accountability_Consultation_14-Oct-13_v3.pdf [Last update unknown] [Last accessed 28th October 2013].

DfE (2013c), *Science GCSE subject content and assessment objectives*, available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/211219/GCSE_Sciences_final_updated.pdf [Last update unknown] [Last accessed 29th October 2013].

Dillon, J (2011) Science Communication—A UK perspective, *International*

Journal of Science Education, Part B: Communication and Public Engagement, 1(1), 5-8.

Ely, M. (1991) *Doing Qualitative Research: Circles within Circles*, London, The Falmer Press.

Francis, B., Hutchings, M. and Read, B. (2004) *Science in Girls' Schools: Factors that contribute to girls' engagement and attainment*, London, London Metropolitan University.

Freebody, P. (2003) *Qualitative Research in Education: Interaction and Practice*, London, Sage.

Funk, H. (2012) Four Models of Language Learning and Acquisition and Their Methodological Implications for Textbook Design, *Electronic Journal of Foreign Language Teaching*, 9(1), 298-311.

Galton, M, Gray, J. and Ruddock, J. (2003) *Transfer and Transitions in the Middle Years of Schooling, 7-14, Continuities and Discontinuities in Learning*, Nottingham, DfES.

Goodman, A. and Gregg, P. (Eds) (2010) *Poorer Children's Educational Attainment: how important are attitudes and behaviour?*, York, Joseph Rowntree Foundation.

Gorard, S., See, B. and Davies, P. (2011) *Do attitude and aspirations matter in education? A review of the research evidence*, Saarbrücken, LAP Lambert.

Gorard, S. (2013) *Research Design: Creating Robust Approaches for the Social Sciences*, London, Sage.

Hargreaves, D.H. (1998) *Social Relations in a Secondary School*, London, Routledge.

Harris, L. (2011) Secondary teachers' s conceptions of student engagement: Engagement in learning or schooling? *Teaching and Teacher Education*, 2(2), 376-386.

Harrison, S. and Morgan, R. (2012) Using Simplified English to identify potential problems for non-native speakers in the language of engineering examination papers, *Language Learning Journal*, 40(1), 113-123.

Hinton, P.R. (1995) *Statistics Explained: A Guide for Social Science Students*, London, Routledge.

Hoyle, P. and Stone, C. (2000) Developing the literate scientist. In: Capel, S., Davidson, J., Arthur, J. and Moss, J. (Eds.) (2000) *Issues in Science Teaching*, London, Routledge/Farmer, 89-99.

Hull, C. (1985) Between the Lines: The analysis of interview data as an exact art, *British Educational Research*, 11(1), 27-32.

IOP (2012) *It's different for Girls: The influence of schools*, London, IOP [available online at: http://www.iop.org/education/teacher/support/girls_physics/file_58196.pdf] [Last update unknown] [Last accessed 24th July 2014].

Knipper, K.J. and Duggan, T.J. (2006) Writing to learn across the curriculum: Tools for comprehension in content area classes, *The Reading Teacher*, 59(5), 462-470.

Krueger, R.A. and Casey, M.A. (2000) *Focus Groups (3rd Ed.) A Practical Guide for Applied Research*, London, Sage.

Lakin, S. and Wellington, J.J. (1994) "Who will teach the nature of Science?" Teachers' views of science and their implications for science education, *International Journal of Science Education*, 16(2), 175-190.

Lara-Alecio, R., Tong, F., Irby, B. J., Guerrero, C., Huerta, M. and Fan, Y. (2012) The effect of an instructional intervention on middle school english learners' science and english reading achievement, *Journal of Research in Teaching Science*, 49(8), 987-1011.

Lee, O. and Fradd, S.H. (1998) Science for All, Including Student From Non-English-Language Backgrounds, *Educational Researcher*, 27(4), 12-21.

PHYSICS KEYWORDS

Levesley, M., Baggley, S., Clarke, J., Gray, S., Johnson, P. and Pimbert, M. (2002a) *Exploring Science 7*, Harlow, Pearson.

Levesley, M., Baggley, S., Clarke, J., Gray, S., Johnson, P. and Pimbert, M. (2002b) *Exploring Science 8*, Harlow, Pearson.

Levesley, M. (Ed.) (2011) *EdExcel GCSE Science*, Pearson, London

Lewis, A. (1992) Group Child Interviews as a Research Tool, *British Educational Research Journal*, 18(4), 413-421.

Little, J.L., Bjork, E.L., Bjork, R.A. and Angello, G. (2012) Multiple-Choice Tests Exonerated, at Least of Some Charges, *Psychological Science*, 23 (11), 1337-1244.

Mason, M., Mason, B. and Quayle, T. (1992) Illuminating English: how explicit language teaching improved public examination results in a comprehensive school, *Educational Studies*, 18(3), 341-353.

McAteer, M. (2013) *Action Research in Education*, London, Sage.

McNiff, J. and Whitehead, J. (2006) *All You Need to Know about Action Research*, London, Sage.

Mertens, D.M. (1998) *Research Methods in Education and Psychology*, London, Sage.

Miles, M.B. (1979) Qualitative Data as Attractive Nuisance: The Problem of Analysis, *Administrative Science Quarterly*, 24(4), 590-601.

Mohan, B. and Slater, T. (2006) Examining the theory/practice relation in a high school science register : A functional linguistic perspective, *Journal of English for Academic Purposes*, 5(4), 302-316.

Mujtaba, T. and Reiss, M.J. (2013) Inequality in Experiences of Physics Education: Secondary School Girls' and Boys' Perceptions of their Physics Education and Intentions to Continue with Physics After the Age of 16, *International Journal of Science Education*, 35(11), 1824-1845.

Nation, P. (2007) The Four Strands, *Innovation in Language Learning and Teaching*, 1(1), 2-13.

Obama, B. (2009) *Remarks by the President in Town Hall Meeting on Health Care*, Washington, The White House, available at:
http://www.whitehouse.gov/the_press_office/Remarks-by-the-President-in-Town-Hall-Meeting-on-Health-Care-in-Green-Bay-Wisconsin/ [Last update unknown] [Last accessed 27th May 2014].

OFSTED (2013) 2012/13 *The Report of Her Majesty's Chief Inspector of Education, Children's Services and Skills: Schools*, Manchester, OFSTED, available at:
<http://www.ofsted.gov.uk/resources/ofsted-annual-report-201213-schools> [Last update unknown] [Last accessed 22nd April 2014].

OFSTED (2014) *Science survey visits*, Manchester, Oxford, available at:

http://www.ofsted.gov.uk/sites/default/files/documents/inspection--forms-and-guides/supplementary-subject-specific-guidance_2/Supplementary%20subject-specific%20guidance%20for%20science.doc [Last update unknown] [Last accessed 22nd April 2014].

Pampaka, M., Williams, J., Hutcheson, G., Wake, G., Black, L., Davis, P. and Hernandez-Martinez, P. (2012) The association between mathematics pedagogy and learners' dispositions for university study, *British Educational Research Journal*, 38(3), 473-496.

Pampush, J. D. and Petto, A. J., (2011) Familiarity with Latin and Greek anatomical terms and course performance in undergraduates, *Anatomical Science Education*, 4(1), 9-15.

Pearson Education Ltd. (2013a), *EdExcel GCSE Physics: Unit P1: Universal Physics Higher Tier*, Wednesday 5th June 2013, London, Pearson.

Pearson Education Ltd. (2013b), *Mark Scheme (Results) Summer 2013 GCSE Physics (5PH1H) Paper 01*, London, Pearson.

Perera, K. (1980) The Assessment of Linguistic Difficulty in Reading Material, *Educational Review*, 32(2), 151-161.

Powney, J. and Watts, M. (1987) *Interviewing in Educational Research*, London, Routledge & Kegan Paul.

Pring, R. (2013) *The Life and Death of Secondary Education for All*, London, Routledge.

QCA (2006) *GCSE Criteria for Science*, available online at:

http://webarchive.nationalarchives.gov.uk/+/http://www.ofqual.gov.uk/files/11881_gcs_e_science_criteria_apr05.pdf [Last update unknown] Last accessed 24th July 2014].

Raz, I. S. and Bryant, P. (1990), Social background, phonological awareness and children's reading. *British Journal of Developmental Psychology*, 8, 209-225.

Reay, D (2006) The Zombie Stalking English Schools: Social Class and Educational Inequality *British Journal of Education Studies*, 54(3), 288-307.

Rose, D. and Sullivan, O. (1996) *Introducing Data Analysis for Social Scientists* (2nd Ed.), Buckingham, OUP.

Ross, K., Lakin, L. and Callaghan, P. (2000) *Teaching secondary science: constructing meanings and developing understanding*, London, Fulton.

Sapsford, R. (1999) *Survey Research*, London, Sage.

Sikes, P. (2000) 'Truth' and 'Lies' Revisited, *British Educational Research Journal*, 26(2), 257-270.

Skribe-Dimec, D. (2013) Differentiation in Teaching Science in the First and Second Cycles of Primary Education, *Journal of Elementary Education*, 6(2-3), 193-211.

Smail, B.(1984) Organising the curriculum to fit girls' interests. In: Kelly, A. (1987) *Science for Girls?*, Milton Keynes, OUP, 80-88.

Stage, E.K., Kreinberg, J.E., Eccles, J. and Becker, J.R. (1985) Increasing the participation and achievement of girls and women in mathematics, science and engineering. In: Kelly, A. (1987) *Science for Girls?*, Milton Keynes, OUP, 119-131.

Svensson, L., Anderberg, E., Alvegard, C. and Johansson, T. (2009) The use of language in understanding subject matter, *Instructional Science*, 37(3), 205-225.

Thomas, G. (2009) *How to do Your Research Project: A guide for students in education and applied social sciences*, London, Sage.

Thompson, S.J. (1999) *Designing Surveys That Work! A step-by-step guide*, London, Sage.

Toplis, R. (2011) Students' Views About Secondary School Science Lessons: The Role of Practical Work, *Research in science education*, 42(3), 531-549.

Traianou, A. (2012) Science teaching: a dilemmatic approach, *Ethnography and Education*, 7(2), 213-226.

Verma, G.K. and Mallick, K. (1999) *Researching Education: Perspectives and Techniques*, London, Falmer Press.

Walford, G. (2001) *Doing Qualitative Educational Research: A personal guide to the research process*, London, Continuum.

Wallace, C. S. (2012) Authoritarian science curriculum standards as barriers to teaching and learning : An interpretation of personal experience, *Science Education*, 96(2), 291-310.

Waring, R. and Takaki, M. (2003) At What Rate Do Learners Learn and Retain New Vocabulary from Reading a Graded Reader?, *Reading in a Foreign Language*, 15(2), 130-163.

Watts, M. and Ebbutt, D. (1988) Sixth formers' views of their science education, 16-18, *International Journal of Science Education*, 10(2), 211-219.

Wellington, J. (2000a) *Educational Research: Contemporary Issues and Practical Approaches*, London, Continuum.

Wellington, J. (2000b) *Teaching and learning secondary science: contemporary issues and practical approaches*, London, Routledge.

Wellington, J. and Ireson, G. (2008) *Science Learning, Science Teaching*, Abingdon, Routledge.

Wellington, J. and Osborne, J.N. (2001) *Language and Literacy in Science Education*, Buckingham, OUP.

Wellington, J. and Osborne, N. (2003) *Language and Literacy in Science Education*, Buckingham, OUP.

Willms, J. (2003) *Student Engagement at School: A sense of belonging and participation. Results from PISA 2000*, OECD report, available at:
<http://www.oecd.org/education/preschoolandschool/programmeforinternationalstudentassessmentpisa/33689437.pdf> [Last update unknown] [Last accessed 1st October 2013].

Zuber-Skerritt, O. (Ed.) (1996) *New Directions in Action Research*, London, Falmer Press.

Appendix A – Physics 1 Keyword List

The Solar System

Orbits
Geocentric
Heliocentric
Telescope
Naked eye
Visible light
Light waves
Radio waves
Microwaves

Refracting Telescopes

Refraction
Interface
Normal
Lens
Converging Lens
Convex Lens
Converge
Focal length
Image
Refracting telescope
Objective lens
Eyepiece lens
Magnify

Lenses

Magnified
Real image
Virtual image

Reflecting telescopes

Reflected
Magnifications
Reflecting telescope
Primary mirror

Waves

Transfer
Transverse waves
Electromagnetic waves
Sound waves
Longitudinal waves
Seismic waves
Frequency
Hertz (Hz)
Wavelength
Amplitude
Wave speed

Beyond the Visible

Infrared
Infrared radiation (IR)
Ultraviolet
Ultraviolet radiation (UV)
Electromagnetic radiation

The Electromagnetic Spectrum

Vacuum
X-rays
Gamma rays
Electromagnetic spectrum
Negative powers

Electromagnetic Dangers

DNA
Skin cancer
Mutations

Using Electromagnetic Radiation

Illuminated
Fluorescence
Fluorescent lamps
Thermal imaging
Optical fibre
Radiotherapy

Ionising Radiation

Ionising radiation
Ions
Radioactive
Alpha (α) particles
Beta (β) particles

The Universe

Stars
Nebulae
Milky Way
Solar System
Galaxy
Universe

Spectrometers

Spectrum
Spectrometer

Exploring the Universe

Visible light

PHYSICS KEYWORDS

Alien Life?

Landers
Space probes
Rovers
Search for Extraterrestrial Intelligence (SETI)

Life-cycles of Stars

Dense
Protostar
Nuclei
Fusion reactions
Main sequence
Red giant
White dwarf
Red supergiants
Supernova
Black hole
Neutron star

Theories about the Universe

Red-shift
Big Bang theory
Steady State theory
Cosmic microwave background (CMB) radiation

Red-shift

Pitch
Doppler Effect
Sound wave

Infrasound

Frequency
Medium
Infrasound

Ultrasound

Ultrasound
Sonar
Ultrasound scan

Seismic waves

Earthquakes
Seismic waves
Seismometers
Focus
Epicentre
P waves
S waves
Crust
Mantle
Core
Refracted

Detecting Earthquakes

Tectonic plates
Convection currents
Tsunami

Renewable Resources for Electricity

Electricity
Current
Voltage
Renewable energy resources
Solar cells
Solar energy
Hydroelectricity
Wind turbines
Geothermal energy
Tidal power
Wave power

PHYSICS KEYWORDS

Non-renewable resources

Non-renewable resources
Fossil fuels
Nuclear power
Climate change
Acid rain
Decommissioned

Generating Electricity

Magnetic field
Electromagnetic induction
Induced current
Dynamamos
Direct current (DC)
Generator
Slip rings
Carbon brushes
Alternating current (AC)
Electromagnets

Transmitting Electricity

National Grid
Efficiency
Transformer
Step-up transformer
Step-down transformer
Primary coil
Secondary coil

Paying for Electricity

Joule (J)
Power
Watt (W)
Kilowatts (kW)
Kilowatt-hour (kWh)
Unit

Reducing Energy Use

Payback time
Cost-efficient

Energy transfers

Thermal
Light
Electrical
Kinetic
Sound
Chemical potential
Nuclear potential
Elastic potential
Gravitational potential
Energy transfer
System
Law of conservation of energy

Appendix B – Words focussed upon in the physics intervention

1	Geocentric	Earth at the centre.
2	Heliocentric	The Sun at the centre.
3	Light waves	The name given to waves in the EM spectrum that we can see.
4	Radio waves	The name given to waves in the EM spectrum that have the longest wavelength.
5	Microwaves	The waves in the EM spectrum that can be used for heating water and sending signals.
6	Interface	The point where two things interact or meet.
7	Real image	An image which can be projected onto a screen.
8	Virtual image	An image which appears to exist if you look through a lens but you cannot project.
9	Transverse waves	Waves where the movement is at right angles to the direction of travel.
10	Electromagnetic waves	"Light waves" found anywhere in the EM spectrum.
11	Sound waves	Waves generated by noise.
12	Longitudinal waves	Waves where the movement is along the direction of travel.
13	Seismic waves	Waves created by earthquakes.
14	Frequency	How many waves pass per second.
15	Wavelength	The length of a wave.
16	Amplitude	The "height" of a wave or how much it is displaced.
17	Infrared	The region of the electromagnetic spectrum that is produced by all objects with heat.
18	Infrared radiation (IR)	The region of the electromagnetic spectrum that is produced by all objects with heat.
19	Ultraviolet	The region of the electromagnetic spectrum that is next to violet but can't be seen.
20	Ultraviolet radiation (UV)	The region of the electromagnetic spectrum that is next to violet but can't be seen.
21	Electromagnetic radiation	Any emitted part of the EM spectrum.
22	Vacuum	A space with no particles in it.
23	X-rays	The region of the spectrum used to take medical images of bones.
24	Gamma rays	The region of the spectrum that kills cells and can be very harmful.
25	Electromagnetic spectrum	The spectrum of light - visible and not visible.
26	Fluorescent lamps	A lamp that gives off light.
27	Thermal imaging	Making an image based on the heat given off by something.
28	Radiotherapy	Treating someone with radiation.
29	Ionising radiation	Radiation that can cause particles to gain a charge by losing electrons or an electron.
30	Ions	A particle with a charge (due to an unbalanced number of electrons).
31	Radioactive	Something which gives off ionising radiation.
32	Alpha (α) particles	Helium nuclei (2 protons and 2 neutrons).
33	Beta (β) particles	High speed electrons.
34	Dense	Tightly packed.
35	Protostar	The beginnings of a new star.
36	Nuclei	The centres of things (including atoms).
37	Main sequence	The normal life cycle a star follows.
38	Red giant	A main sequence star that is getting older and bigger.
39	White dwarf	The remnants of a dead main sequence star.
40	Red supergiants	A larger star that is getting older and bigger.
41	Supernova	A huge explosion from a dying star.
42	Black hole	A dead star creating so much gravitational pull that even light is pulled in.

PHYSICS KEYWORDS

43	Neutron star	A dead star that is very dense.
44	Big Bang theory	The idea that the Universe quickly expanded from a point in space and no new matter is created.
45	Steady State theory	The idea that the Universe has always existed but gets bigger with new matter being made.
46	Cosmic microwave background (CMB) radiation	Radiation found evenly spread throughout the Universe.
47	Pitch	Effectively means frequency.
48	Doppler Effect	The effect heard when an ambulance comes towards you then moves away.
49	Sound wave	A wave caused by noise.
50	Medium	A material through which something passes.
51	Infrasound	Low frequency sound (below 20Hz).
52	Ultrasound	High frequency sound (above 20,000Hz).
53	Sonar	A system used to see how deep water is (or look for objects in water).
54	Seismic waves	A wave which travels through the Earth.
55	Focus	The point where a seismic wave began.
56	Epicentre	The point at the surface directly above the focal point of an earthquake.
57	P waves	Longitudinal (and faster) seismic waves.
58	S waves	Transverse (and slower) seismic waves.
59	Tectonic plates	The large area of the Earth's crust that move slowly.
60	Convection currents	The way in which heat "moves" in the mantle, rising and falling.
61	Tsunami	Water waves caused by movement in the Earth's crust.
62	Electricity	The flow of charge/electrons, often in a wire.
63	Current	The flow of charge/electrons.
64	Voltage	The "pressure" in a circuit, sometimes called the potential difference.
65	Renewable energy resources	Energy that comes from a source that is continually replenished (isn't running out).
66	Solar energy	Energy that comes from sunlight.
67	Hydroelectricity	Electricity produced by water turbines in dams or reservoirs.
68	Geothermal energy	Energy from hot rocks beneath the Earth's surface.
69	Tidal power	Energy from the rising and falling tide.
70	Wave power	Energy from waves.
71	Non-renewable resources	Energy resources from sources that cannot be replenished (can run out).
72	Nuclear power	Energy from nuclear materials like uranium.
73	Acid rain	Rain with a lower pH.
74	Magnetic field	A field caused by a magnet with a north and south.
75	Induced current	Current created by a magnetic field being moved along a wire (like in a dynamo).
76	Dynamos	A device which used movement (kinetic energy) to create electricity.
77	Direct current (DC)	Current which flows in one direction only.
78	Generator	A device which produces electricity.
79	Slip rings	Connection rings found on a dynamo.
80	Carbon brushes	Connectors used to maintain a good contact with the dynamo as it spins.
81	Alternating current (AC)	Current which flows "back and forth" or in different directions.
82	Electromagnets	A magnet created using an electrical current.
83	National Grid	The wires and transformers that supply electricity around the country.
84	Transformer	A device which changes the voltage of AC current.
85	Step-up transformer	A device which increases the voltage of AC current.
86	Step-down transformer	A device which decreases the voltage of AC current.
87	Primary coil	The coil of wire carrying electrical current entering a transformer.

PHYSICS KEYWORDS

88	Secondary coil	The coil of wire carrying electrical current leaving a transformer.
89	Thermal	Heat energy.
90	Electrical	Energy which is the flow of electrons.
91	Kinetic	Movement energy.
92	Sound	Energy we hear.
93	Chemical potential	Energy potential in chemicals.
94	Nuclear potential	Energy potential in radioactive materials.
95	Elastic potential	Energy potential in something elasticated.
96	Gravitational potential	Energy potential as you move up through gravity or higher.
97	System	Connected processes that make up a bigger whole.
98	Law of conservation of energy	The law that energy is never lost.
99	Converging Lens	A lens which focuses light inwards onto a point. Also called convex.
100	Convex Lens	A lens which focuses light inwards onto a point. Also called converging.

Words highlighted in yellow were used in the word test.

Appendix C – Ethical Procedure and Consents

- Research information sheet for participants
- Parent consent opt-in
- Teacher consent opt-in

Research Information Sheet for Participants

This sheet is designed to explain the research in which you have agreed to participate.

The research question being studied is:

“Are there discernible benefits to pupils in Physics when science pedagogy is developed to include more vocabulary-based development skills for the pupils to make use of during lessons.”

This leads to the following sub-questions:

1. Do pupils show an improved ability to define keywords following the use of a pedagogical intervention in during a GCSE Physics module.
2. What worthwhile benefits do pupils perceive from the change in pedagogy?
3. What worthwhile benefits do teachers of the pupils perceive from the change to pedagogy?

Stage 1

A set of sheets will be used during Physics lessons which encourage pupils to use input and output methods in reading, writing, speaking and listening to improve their understanding of around 100 keywords in the Physics 1 module of GCSE Core Science/Physics.

Pupils will be tested before and after the module to check how much additional understanding has been gained. The data from this will be used anonymously to analyse the benefits statistically.

Stage 2

Focus groups will be carried out with pupils and teachers of the module to try to understand what benefits there are from this approach. The groups will be recorded and analysed later. All responses will be anonymous.

The Thesis

The data collected will be used to analyse the questions above. This will be recorded as a thesis of up to 40,000 words. This thesis will be submitted to the University of Birmingham as part of a MA in Research (Education).

A copy of the thesis will be kept in the University library. Another copy will be issued to school.

Use of the findings

Science staff in school will use the findings to reflect upon the needs of pupils in school. This may involve sharing these findings with other staff in school.

The research could potentially lead to more studies.


Anonymity

No results or comments in the thesis will be able to be linked to individuals. Names will not be given to anyone outside of school.

Data Protection

Data will be password protected and held securely.

All research will be carried out in accordance with the laws of the United Kingdom, the ethics policies of the University of Birmingham and the policies of the school.



PHYSICS KEYWORDS

[School header removed]



PHYSICS KEYWORDS



PHYSICS KEYWORDS

[School header removed]

Dear <teacher>,

As you will be aware, I am currently conducting post-graduate research at the University of Birmingham.

As part of this research into the development of pupil understanding of keywords, I would like to conduct a focus group with the science staff. As such I hope that you would like to be involved.

The data gathered from this focus group with of course be used anonymously with identities kept confidential outside of the focus group.

You may choose not to participate in this focus group without any further recourse or impact. You may also request that your contribution is not used in the research within 14 days of the focus group.

Please do let me know if you would like to discuss the research further. I will keep the Science Blog updated as the research progresses.

I am willing to participate in a focus group as part of research with the University of Birmingham.

Name: _____

Signature: _____ Date: _____

Appendix D – Example of a Keyword Sheet

Topic: The Solar System

Listen

Listen to your teacher explain what each of these words means then work in pairs to write down a meaning in each box.

Geocentric	
Heliocentric	
Convex Lens	

Read

Read the paragraph below. Then discuss with your partner whether the definitions you wrote above are correct.

People used to believe that the solar system was **geocentric**. This means they thought that the Earth was in the middle and that the Sun, planets and stars moved around the Earth. We now know that the solar system is **heliocentric**. This means that the Sun is at the centre and the planets orbit the Sun. We have also learnt that some things orbit each planet, like moons. We were able to learn this because people managed to make better and better **convex lenses** to use in telescopes so that they could magnify smaller objects and let us see even further into space.

Speak

Imagine that you have just discovered that the Sun is at the centre of the solar system. All of your friends believe the sun is at the centre. Tell your partner what you have found out, imagining they are one of your friends who does not know this news! Try to make it interesting and exciting!

Discuss with each other how you got on when you have both had a go!

Write

Write down your short speech using your practice and feedback.

--

Develop (Extension)

Explain whether the Universe (all of the stars and galaxies in space) can be thought of as heliocentric. You can do this as a discussion or write down your thoughts.

Appendix E – Word Test

Physics 1 – Key word understanding

Complete all questions as well as you can. There is no strict time limit but this should take around 30 minutes. Do not talk, ask for help or copy.

Circle the meaning which is the closest meaning to the key word in bold. If you don't know then circle "I don't know".

Geocentric	Sun at the centre.	Earth at the centre.	A mathematical shape.	Geography based.	I don't know
Light waves	Waves that don't weigh much.	Someone waving so that you can see them.	The name given to waves in the EM spectrum that we can see.	Waves you can hear.	I don't know
Microwaves	The waves in the EM spectrum that can be used for heating water and sending signals.	Waves made by a microprocessor.	The smallest EM waves.	Waves that can carry sound in the sea.	I don't know
Real image	An image which we can't actually see.	An image of a Spanish football club.	An image which can be projected onto a screen.	A shadow.	I don't know
Transverse waves	A wave where the energy travels in the direction of the displacement.	A wave from a transformer.	Sound waves.	Waves where the movement is at right angles to the direction of travel.	I don't know
Sound waves	Waves that come from the Sun.	Waves generated by noise.	Waves that travel in space.	Waves that mobile phones send to satellites.	I don't know
Seismic waves	Waves created by earthquakes.	Waves created by size.	Waves created by sound.	Waves created by a computer.	I don't know
Wavelength	The frequency of a wave.	The length of a wave.	The amplitude of a wave.	The speed of a wave.	I don't know
Infrared	The region of the electromagnetic spectrum that is next violet.	The region of the electromagnetic spectrum that is smallest.	The region of the electromagnetic spectrum that is next biggest.	The region of the electromagnetic spectrum that is produced by all objects with heat.	I don't know
Ultraviolet	The region of the electromagnetic spectrum that is produced by all objects with heat.	The region of the electromagnetic spectrum that is smallest.	The region of the electromagnetic spectrum that is next to violet but can't be seen.	The region of the electromagnetic spectrum that is next biggest.	I don't know

PHYSICS KEYWORDS

Electromagnetic radiation	A helium nucleus.	A high speed electron.	Energy created by a big magnet.	Any emitted part of the EM spectrum.	I don't know
X-rays	Waves used to send radio signals.	Heat waves.	Sound waves.	The region of the spectrum used to take medical images of bones.	I don't know
Electromagnetic spectrum	The visible part of the EM spectrum.	A rainbow.	A 1980s computer.	The spectrum of light - visible and not visible.	I don't know
Thermal imaging	A camera negative.	Making an image based on the heat given off by something.	Using a hot knife to make an image.	A type of insulating blanket.	I don't know
Ionising radiation	Radiation that creases clothes.	Visible light.	Radiation that can cause particles to gain a charge by losing electrons or an electron.	Radiation that can cause particle to become isotopes.	I don't know
Radioactive	A radio that can be used for sport.	A radio that is turned on.	A bomb.	Something which gives off ionising radiation.	I don't know
Beta (β) particles	A helium nucleus.	High speed electrons.	An EM ray.	A gamma ray.	I don't know
Protostar	A dead star.	A pulsar.	A black hole.	The beginnings of a new star.	I don't know
Main sequence star	A very large star that may form a black hole when it dies.	The normal life cycle a star follows.	The centre of a fusion torus.	A pulsar.	I don't know
White dwarf	The remnants of a dead main sequence star.	A small man.	A small galaxy.	A small moon.	I don't know
Supernova	A type of champagne.	A huge explosion from a dying star.	A black hole.	A nebula.	I don't know
Neutron star	A very large red star.	A very large white star.	A dead star that is very dense.	A blue star.	I don't know
Steady State theory	The idea that the Universe was created in a huge explosion.	The idea that the Universe has always existed but gets bigger with new matter being made.	The idea that the Universe is always in the same state.	A religious idea about the Universe.	I don't know
Pitch	Amplitude.	Effectively means frequency.	Loudness.	Wavelength.	I don't know

PHYSICS KEYWORDS

Sound wave	A wave caused by noise.	A wave caused by light.	A wave found in space.	A P wave.	I don't know
Infrasound	High frequency sound (above 20,000Hz)	Low frequency sound (below 20Hz).	Sound that can be heard by humans.	Light waves.	I don't know
Sonar	A way of measuring how big the sun is.	A type of note on a musical instrument.	A sound made during an earthquake.	A system used to see how deep water is (or look for objects in water).	I don't know
Focus	The nearest town to an earthquake.	The place on the surface above the point where an earthquake starts.	The place where P and S waves are furthest apart.	The point where a seismic wave began.	I don't know
P waves	The slowest waves from an earthquake.	Transverse waves from an earthquake.	Secondary waves.	Longitudinal (and faster) seismic waves.	I don't know
Tectonic plates	Plates built around a nuclear power station to protect it.	The large area of the Earth's crust that move slowly.	Plates used to generate sound in a loudspeaker.	Plates used to dissect atoms on.	I don't know
Tsunami	A large hurricane.	Water waves caused by movement in the Earth's crust.	A huge explosion.	Sound waves caused by an explosion.	I don't know
Electrical current	The flow of charge/electrons.	The power of a circuit.	The voltage of a circuit.	The built up of charge in an insulator.	I don't know
Renewable energy resources	Energy from fossil fuels.	Energy that comes from a source that is continually replenished (isn't running out).	Energy from nuclear fission power stations.	Energy from electric cars.	I don't know
Hydroelectricity	Electricity produced by the tide.	Electricity produced by a nuclear power station.	Energy produced by solar panels.	Electricity produced by water turbines in dams or reservoirs.	I don't know
Tidal power	Energy from a dam on a river.	Energy from the rising and falling tide.	Energy from solar panels.	Energy from inland lakes.	I don't know
Non-renewable resources	Fossil fuels (coal, oil and gas).	Solar power.	Wind turbines.	Energy resources from sources that cannot be replenished (can run out).	I don't know
Acid rain	Rain with a high pH (above 8)	Rain with a pH of 7.	An acid spillage in the lab.	Rain with a lower pH.	I don't know
Induced current	Static charge.	Current created by a magnetic field being moved along	Current created by a mixing chemicals.	Current being changed to a lower level.	I don't know

PHYSICS KEYWORDS

		a wire (like in a dynamo).			
Direct current (DC)	Current which flows in one direction only.	Current which flows in both ways around a circuit.	Current which does not flow.	Current which flows only in straight lines.	I don't know
Slip rings	A ring people wear to reduce static build up.	A ring which slips over a bunsen burner.	A type of bell.	Connection rings found on a dynamo.	I don't know
Alternating current (AC)	Current which flows "back and forth" or in different directions.	Current which flows in one direction around a circuit.	Static electricity.	Current which is very low.	I don't know
National Grid	A large map that is split up into squares to show where everything is.	The wires and transformers that supply electricity around the country.	The railway lines around the country.	A large cover in a road.	I don't know
Step-up transformer	A device which decreases the voltage of AC current.	A device which increases the voltage of AC current.	A device which decreases the voltage of DC current.	A device which increases the voltage of DC current.	I don't know
Primary coil	The coil on a transformer that is the output.	The biggest coil on a transformer.	The middle of an electromagnet.	The coil of wire carrying electrical current entering a transformer.	I don't know
Thermal	Chemical energy.	Movement energy.	Light energy.	Heat energy.	I don't know
Kinetic	Heat energy.	Chemical energy.	Movement energy.	Light energy.	I don't know
Chemical potential	The energy in an elastic band.	The strength of a chemical.	The concentration of a chemical.	Energy potential in chemicals.	I don't know
Elastic potential	The potential to change into something new.	Energy potential in something elasticated.	The energy stored in a chemical battery.	The energy found in a piece of metal.	I don't know
System	Disconnected pieces that do not affect each other.	A piece of a toilet that stores water.	A rule.	Connected processes that make up a bigger whole.	I don't know
Converging Lens	A lens which focuses light outwards onto a point.	A lens which focuses light inwards onto a point. Also called convex.	A lens which does not affect light.	A lens which is very large.	I don't know

You have finished!