An Evaluation of Secondary School Students' Learning Experiences with Astronomy-based Physics Outreach Activities

Sophie Bartlett

A thesis submitted for the degree for Doctor of Philosophy

December 2018

Abstract

England and Wales, as with many other countries across the world are faced with a large proportion of secondary school students who are disengaged with physics. Teachers are faced with immense pressure from accountability measures that have been seen to restrict innovation and creativity in the classroom. This study implemented a series of astronomy-based outreach activities that had been designed to apply a novel context to the curriculum and encourage an inquiry-based, student-centred classroom. Though astronomy has been described as a promising point of engagement and portrays widespread interest, no study is seen to evaluate the application of astronomy as a context for compulsory science or physics curricula. This study successfully achieves this through the implementation of a series of outreach activities that apply an inquiry-based pedagogy. The study focuses on the evaluation of their impact on students' learning experiences in physics. However, findings are not limited to the outreach activities of this study and a robust framework is presented to other practitioners seeking to design and evaluate their outreach initiatives and teachers looking to implement them.

The study also demonstrates that initiatives do not necessarily require large amounts of time or funding in order to have an impact. Though existing research is indicative that attitudes are withstanding and difficult to change, this study demonstrates the influence just a single lesson can have on students' perceptions of their learning experiences and can encourage them to reconsider their predispositions.

Table of Contents

TABLE OF CONTENTS	
LIST OF TABLES AND FIGURES	VIII
ACKNOWLEDGEMENTS	X
1. OVERVIEW AND PURPOSE OF STUDY	1
2. SCIENCE AND PHYSICS EDUCATION IN WALES AND ENGLAND	4
2.1 Education in England and Wales	4
2.1.1 School Accountability	5
2.2 Applicants in Science and Physics	7
2.5 The Importance of STEM and Demand for its Professions	10
2.4 Previous Science and Physics Oureach Initiatives 2.5 Identifying the Point of Disengagement	11
2.5 Identifying the Folint of Disengagement 2.5.1 Who Stays and Who Leaves?	17
2.5.1 Who stays and who Leaves: 2.5.2 Summary	20
2.6 Secondary School Students' Perceptions of Science and Physics	20
2.6.1 Interest and Enjoyment	21
2.6.2 Confidence versus Actual Ability	23
2.6.3 Relevance	24
2.6.4 Usefulness and Career Aspirations	26
2.6.5 Classroom Experiences	27
2.6.6 Summary	30
2.7 Aims of the Study	31
3. INSTRUCTIONAL DESIGN	33
3.1 Learning Theories	33
3.2 Inquiry-based Learning	36
3.2.1 The Case for Inquiry-based Learning	37
3.2.2 Optimum Level of Inquiry-based Learning	39
3.3 Limitations of Inquiry-based Learning	41
3.4 The Inquiry Approach for this Study	42
3.5 Summary	44
4. PHYSICS-BASED OUTREACH ACTIVITIES	45
4.1 Contextualising Physics	45
4.2 Robotic Telescopes	47
4.2.1 Robotic Telescopes in Science Education	48
4.2.2 The Faulkes Telescope Project	52
4.3 Down2Earth	54
4.3.1 Computer Simulations in Science Education	54
4.4 Educational Outreach Activities	58
4.4.1 KS5: The Impact Calculator	59
4.4.2 KS5: Hubble's Expansion of the Universe	60 41
4.5 Summary	01

5. I	EVALUATION FRAMEWORK	62
5.1	The Concept of Interest	65

5.2 Self-Determination Theory	66
5.3 Individuals' Internal Goals	67
5.4 The Teacher's Role in the Classroom	68
5.5 Constructing the Framework	68
5.5.1 Goal Orientation	71
5.5.2 Differentiation	71
5.5.3 Student Autonomy	72
554 Collaborative/Cooperative Learning	72
5.5.5 Context Novelty	73
5.5.6 Teacher Role	74
5.5.7 Determining Overall Effectiveness	75
5.6 Summary	75
5.6 Summary	15
6. METHODOLOGICAL DESIGN	77
6.1 The Researcher: Predispositions, Beliefs and Biases	77
6.2 The Nature of Research	80
6.2.2 Summary	83
6.3 Research Questions	84
6.4 Research Design	84
6.4.1 Mixed Method Approach	85
6.4.2 Sequential Explanatory Design	86
6.4.3 Case Studies	87
6.5 Recruiting Schools Teachers and Students	89
6.6 Ethical Procedures	90
6.7 Case Study 1: Down?Farth	92
6.7.1 Group Pink	92
6.7.2 Group Green	93
6.7.3 Group Blue	03
6.7.4 Group Neur	93
6.7.5 Crown Drown	94
0.7.5 Group Brown	94
6.8 Case Study 2: Hubble's Expansion of the Universe	95
6.8.1 Group Yellow	95
6.8.2 Group Orange	96
6.8.3 Group Red	97
6.8.4 Group Black	97
6.8.5 Group Purple	97
6.8.6 Group Grey	98
6.8.7 Summary	98
6.9 Summary	100
7. INSTRUMENTATION FOR DATA COLLECTION	101
7.1 Measuring Perception	101
7.2 Perceptions of Physics Questionnaire	104
7.2.1 Types of Ouestions	105
7.2.2 Designing the PPO	107
7.2.3 Trialling and Refining the PPO	111
7.3 Classroom Observations	114
7.3.1 Developing the Observation Protocol	116
7.4 Student Focus Groups	110
7.5 Teacher Interviews	121
7.6 Summary	123
7.0 Summary	124
8 ADDRESSING VALIDITY AND DELLADILITY	176
0. ADDREDSHNG VALIDITT AND KELIADILITY 2.1 Cardibility	120
8.1 Credibility	127
8.2 Confirmating	129
8.3 I ransierability	130

V

	VI
8.4 Consistency	132
8.5 Summary	133
9. PRELIMINARY DATA ANALYSIS	134
9.1 Quantitative Data Analysis	134
9.1.1 Case Study 1: Down2Earth	135
9.1.2 Case Study 2: Hubble's Expansion of the Universe	139
9.2 Qualitative Data Analysis	143
9.2.1 Thematic Analysis of Observational Data	144
9.2.2 Content Analysis of PPQ Open Questions	145
9.3 Summary	147
10. CASE STUDY 1: DOWN2EARTH	148
10.1 Goal Orientation	148
10.2 Differentiation	151
10.3 Student Autonomy	157
10.4 Collaborative/Cooperative Learning	162
10.5 Context Novelty	164
10.6 Teacher Role	167
10.7 Overall Effectiveness	171
11. CASE STUDY 2: HUBBLE'S EXPANSION OF THE UNIVERSE	173
11.1 Goal Orientation	173
11.2 Differentiation	175
11.3 Student Autonomy	181
11.4 Collaborative/Cooperative Learning	183
11.5 Context Novelty	186
11.6 Teacher Role	189
11.7 Overall Effectiveness	192
12. DISCUSSION	194
12.1 Research Question 1	194
12.1.1 Investigation and Exploration	195
12.1.2 Experiences of Autonomy	196
12.1.3 Cooperation and Collaboration	197
12.1.4 Novel, Unexpected Experiences	197
12.1.5 Effective Differentiation for Building Confidence	197
12.2 Research Question 2	198
12.3 Research Question 3	199
12.3.1 Opportunity: Overcomes Barriers to Reform	200
12.3.2 Opportunity: An Unexpected, Novel Point of Engagement	200
12.3.3 Opportunity: Outreach Activities Preventive of Entirely Didactic Teaching	201
12.3.4 Limitation: Differentiation Requires Improvement	201
12.3.5 Limitation: The Role of the Teacher 12.4 Research Question 4	201 202
13. CONCLUSION	205
14. REFERENCES	210
15. APPENDICES	231
Annendix 1 – Down?Earth 'Recreate a Crater' Resource	231
Appendix 2 – Hubble's Expansion of the Universe Resource Extract	231
	255

Appendix 3 – Methods of Recruiting Schools and Teachers	241
Appendix 4 – Invitation Email to Teachers	244
Appendix 5 – Ethical Approval Form	245
Appendix 6 – Ethical Approval	248
Appendix 7 – Information Sheets and Consent Forms	249
Appendix 8 – Students' Perceptions of Physics Questionnaire (Pre)	252
Appendix 9 – Students Perceptions of Physics Questionnaire (Post)	254
Appendix 10 – Classroom Observation Demographics and Protocol	256
Appendix 11 – Student Focus Groups	259
Appendix 12 – Teacher Interview Questions	260
Appendix 13 – Examples of Thematic Analysis of Observational Data	262
Appendix 14 – Examples of Content Analysis of Open PPQ Responses	265
Appendix 15 - Students Responses to what they Liked Most about the D2E Activity relating to Context 2	273

List of Tables and Figures

Table

Page

		0
Table 1	Positive and Negative Experiences in Physics Education	21
Table 2	Variations in Levels of Inquiry	39
Table 3	How Structured Inquiry Tackles Barriers to Constructivism	44
Table 4	Results from Sarabando et al. (2014, p.119)	56
Table 5	Key of Evaluation Framework Mind Map	63
Table 6	Existing Evaluation Frameworks within the Literature	70
Table 7	Summary of Participating Groups	99
Table 8	Data Collection Instruments and Implementation Phase	103
Table 9	Number of Items Contributing to each Construct in the PPQ	111
Table 10	Summary of Data Collection Instruments	125
Table 11	Summary of Techniques to Improve Validity and Reliability	133
Table 12	Pre- and Post-PPQ Scale Scores for all Students across both Cases	135
Table 13	Pre- and Post-PPQ Scale Scores for Students in the Down2Earth Case Study	136
Table 14	Skewness Values of D2E Student Scores on Pre- and Post-PPQ	139
Table 15	Pre- and Post-PPQ Scale Scores for Students in the Hubble Case Study	140
Table 16	Skewness Values of Hubble Student Scores on Pre- and Post-PPQ	143
Table 17	Thematic Analysis of Observational Data	145
Table 18	Content Analysis of Open PPQ Responses	146
Table 19	Respondents who Disagreed or Strongly Disagreed with Confidence Scale Items	150
Table 20	D2E Case Study Group Scores on Pre-PPQ General Interest Scale	169
Table 21	Hubble Case Study Students' Post-PPQ Responses indicative of Experiences of Difficulty	176
Table 22	Successful Recruitment Methods for Individual Participant Groups	243
Table 23	Examples of Thematic Analysis of Observational Data	262
Table 24	What do you like <i>most</i> about physics lessons?	265
Table 25	What do you like <i>least</i> about physics lessons?	266
Table 26	What did you like most about the activity? (Down2Earth)	267
Table 27	What did you like <i>least</i> about the activity? (Down2Earth)	268
Table 28	How do you think the activity could be improved? (Down2Earth)	269
Table 29	What did you like <i>most</i> about the activity? (Hubble)	270
Table 30	What did you like <i>least</i> about the activity? (Hubble)	271
Table 31	How do you think the activity could be improved? (Hubble)	272

T T T
· •
Δ
_

_	
Pa	ge

Figure		Page
Figure 1	Percentage of Total A-level Students in England and Wales who sat Exams in Individual Science between 2010 and 2018	7
Figure 2	All Applicants and Physical Science Applicants to University in England and Wales since 2010	9
Figure 3	Students' Preferences of Inquiry. Summary of Results from Chatterjee <i>el al.</i> (2009)	40
Figure 4	A Light Curve of Asteroid (1676) Kariba Created by School Students from Five Different Countries	53
Figure 5	Mind-map Used to Conceptualise the Evaluation Framework	64
Figure 6	Flow Chart of the Review Process of Classroom Observations for each Case	120
Figure 7a	Distributions of D2E Student Scores on Individual PPQ Scales on Pre-test	137
Figure 7b	Distributions of D2E Student Scores on Individual PPQ Scales on Post-test	138
Figure 8a	Distributions of Hubble Case Study Student Scores on Individual PPQ Scales on Pre-test	141
Figure 8b	Distributions of Hubble Case Study Student Scores on Individual PPQ Scales on Post-test	142
Figure 9	Process of Data Analysis	144
Figure 10	Student Responses to Confidence Scale on Post-PPQ in the D2E Case Study	149
Figure 11	Histograms of the Distributions of Student Scores on the Confidence Scale on Pre and Post D2E Implementation	153
Figure 12	Illustration of the Overall Effectiveness of the Down2Earth OAs	172
Figure 13	Example of the Resulting Hubble Plot Produced in the Hubble Activity	177
Figure 14	Distribution of Hubble Case Study Student Scores on Confidence Scale Pre- and Post-Implementation	178
Figure 15	Illustration of the Overall Effectiveness of the Hubble's Expansion of the Universe OAs	193

Acknowledgements

I would like to begin by dedicating this thesis and biggest achievement of my life to my loving grandparents, David and Olga Devall. You two are an inspiration and I cannot thank you enough for your unconditional love and support.

I must express my immense gratitude to Dr. Martin Dill Faulkes. He not only provided the sponsorship to make this PhD possible, but he also gave his support and belief in me which I value just as greatly.

I don't think I will ever be able to thank James Snook enough for his continuous support both academically and emotionally. He has gone above and beyond in helping me, his patience and investment were second to none and every single email and conversation were so appreciated.

Thank you to Dr. Paul Roche for believing in me and providing me with the opportunity not only to complete this PhD but to experience such incredible opportunities that have vastly improved my ability to multitask! He supported me from undergraduate to postgraduate and I would not be where I am today without him.

My everlasting thanks to Dr. Michael Fitzgerald and Professor David McKinnon. You two were the light in a very dark tunnel and honestly, you got me through the hardest part. You introduced me to the wonderful world of astronomy education and your belief in me helped me to believe in myself. David and Olga you are my grandparents from down-under.

My incredible parents and brother, Jack. I love you infinitely and your unconditional love is overwhelming. Thank you for supporting me, trusting me and believing in me even when I didn't believe in myself. I hope this achievement makes you just half of proud as I am of all three of you.

Thank you to my wonderful Grampy, Alex Bartlett. You always make me laugh no matter what and you put a smile on my face whenever I see you. I'm not sure you know what a PhD is but you were there all the same!

To Sarah and Fraser, you have been my sanity checks and have kept me grounded through this entire experience.

To Tyler and my incredible friends, particularly Lori, Ellie, Aoife and Elain. You all knew when to ask questions and when not to. You were always a listening ear when I needed to vent and always had a bottle of wine ready when I needed it.

My godmother Joan, it was so wonderful to hear your stories of academia and you are always so supportive. I'm extremely lucky to have you as my godmother.

Last but by no means least, thank you to Nicola. You provided invaluable support entirely from the goodness of your heart and I simply do not know how to thank you for that.

Chapter 1

Overview and Purpose of Study

Physics is consistently seen to be the least popular science among secondary school students throughout many countries around the world and England and Wales are no exception (Bartlett, 2018; Mujtaba and Reiss, 2013; Barmby and Defty, 2006). This is evident from applicants into physics-based courses post-16 education, literature around students' attitudes and perceptions and concerns expressed by physics-based businesses.

Research is largely indicative that students' attitudes towards physics and science go into decline as they progress through secondary education (Barmby *et al.*, 2008; Osborne and Dillon, 2008; Tymms *et al.*, 2008). A number of themes are embedded in the literature in terms of specific aspects guiding students' negative attitudes in science, including; general disinterest or lack or enjoyment of physics, low confidence in their ability, a lack of perceived relevance of physics to their day-to-day lives, its usefulness to future careers and poor pedagogical experiences. Many of such problems are considered largely related to performance focused accountability measures faced by teachers that sadly encourages teaching-to-test and teacher-centred classrooms.

Such problems are also not new ones and are cited throughout the literature spanning decades. Various initiatives, demonstrating various levels of success, have been developed and implemented across schools in the UK and overseas in an attempt to tackle some of such problems. Although, what is clear from review of the research is the lack of high quality evaluation and evidence of impact (Vennix *et al.*, 2018; Bell, 2014).

Astronomy has been described as the "gateway science" (Salimpour et al., 2018) and an effective "point of engagement" (Osborne and Collins, 2001) due to its natural ability to engage people, regardless of their interest in physics or science more generally. However despite such promises, no formal research has been carried out on the effectiveness of its application as a context to compulsory content of science of physics curricula. This study focuses on the evaluation of a series of physics outreach activities (OAs) that have been developed to apply a context of astronomy to the compulsory physics curriculum in England and Wales.

The OAs are evaluated in terms of their influence on students' perceptions of their learning experiences and students' and teachers' interaction with them. This is not specific or limited to the particular OAs themselves, but in order to identify the components that constitute a 'good and effective OA'. This is defined in terms of its influence on students' learning experiences in

physics and also how they support teachers and do not threaten a conflict with accountability measures. A robust framework for evaluating OAs is developed and a series of recommendations are provided to both to OA developers and practitioners, and for teachers. This study is therefore considered to provide a wider applicability and offer insight to the wider field of research in physics and science education.

This study follows a pre- and post-test mixed method approach with a sequential explanatory, multiple-case study design. (Creswell, 2003; Johnson and Onwuegbuzie, 2004; Teddlie and Tashakkori, 2009; Cohen *et al.*, 2011).

The research is rooted within a constructivist epistemology with a realistic ontological position in that it is the researcher's belief that reality is subjective and dependent on experience, context, and the perspectives of the participants themselves (Creswell and Plano-Clark, 2011). Due to this position, the researcher is very much involved and immersed in the research process, the implementation and the participants, in order to understand their environment and perspectives.

The study does not have a pre-defined hypothesis and therefore data is gathered inductively with goal-free evaluation (Scriven, 1991). Although the study sets out specific research questions to address, the researcher will consider all data that is yielded and will not disregard any evidence that comes to light, regardless of its relevance to the specific research questions. In terms of analysis, interpretations are constructed around reflexivity of the researcher's experiences and understandings of the educational environment and phenomena as it evolves with the study's progress (Cohen *et al.*, 2011).

In order to manage and prioritise such large amounts of data, the researcher follows recommendations from Stake (1995) and develops a concise evaluation framework that structures and refines analysis and results. This framework is also used to synthesise quantitative and qualitative data for analysis.

Context is a consistent theme throughout this study. Not only do the educational OAs provide an astronomy context to the physics curriculum, but the research itself is considered to be contextual. That is, school, classes, students and teachers are not uniform and each classroom has its own internal climate based on infinite variables. Details of each environment and scenario of implementation are therefore provided in order to offer a context to the reader and to enhance the study's transferability.

This thesis is structured around a following twelve chapters. Chapter 2 lays the foundations of the study, discussing the importance of science and physics and where secondary education in such subjects are falling short and struggling to inspire and engage students. Chapter 3 explores the pedagogical approaches to learning and reviews such approaches in order to identify that most suited to the objectives and focus of this study. Chapter 4 brings together the previous two

chapters to explain how the outreach activities implemented in this study are designed to build upon the strengths and counter the limitations of previous initiatives. This chapter also contextualises the outreach activities for the reader, setting the scene for implementation.

Chapter 5 defines the theoretical and evaluation framework that guides the data analysis and interpretations. It defines the theoretical underpinnings of how student perception, interest and engagement is defined and provides a framework to structure and synthesise multiple sources of data and highlight its transferability.

Chapter 6 to 8 specify the methodological approach and design of the study. More specifically, Chapter 6 clarifies the research paradigm in terms of the researcher's epistemological and ontological predispositions and approach to research. Chapter 7 describes the instrumentation implemented for data collection and Chapter 8 highlights the approaches employed to address issues of validity and reliability.

Chapters 9 to 11 report on the data analysis. Chapter 9 explores the preliminary data analysis, providing the 'big picture' and a surface view of the study's findings around the evaluation of the outreach activities. Chapters 10 and 11 then explore the data in greater depth. The multiple evidence sources are triangulated through the lens of the evaluation framework and findings and interpretations are contextualised through a case-by-base analysis.

Chapters 12 and 13 then discuss the overall findings of the study, explaining specifically how they address the initial research questions, why such findings are important and how they contribute new knowledge and understanding to the field of physics education.

Chapter 2

Science and Physics Education in Wales and England

2.1 Education in England and Wales

In the UK, education policies are distinct and independent for each nation, though similarities are evident between England and Wales (Brill *et al.*, 2018). Education in England and Wales is compulsory up until the age of 16, however in England you must choose one of three options to continue with until the age of 18. These are:

- Remain in full-time education, e.g. a sixth form or college;
- Begin an apprenticeship or traineeship;
- Spend 20 hours or more per week working or volunteering, whilst in part-time education or training.

Science is also compulsory in education up to the age of 16 and is considered to be a core subject alongside English (or Welsh) language and mathematics (Wellcome Trust, 2015).

Education standards in England and Wales are regulated by Ofsted, (the Office for Standards in Education), and Estyn, (which translates from the Welsh language to "to reach or extend") respectively. These bodies are responsible for inspecting schools and learning centres to ensure they are of a certain standard. Institutes are classified into one of four possible ratings, in England these are *inadequate*, *requires improvement*, *good* and *outstanding*, and in Wales: *unsatisfactory*, *adequate*, *good* and *excellent*. They also produce subject-specific reports summarising findings from inspections across a representative sample of schools.

In their outline of the national curriculum for science in England, the Department for Education (2015) highlight the purpose of science in education and why it is important. They affirm that students must be taught essential aspects of knowledge, methods, processes and uses of science, emphasising how science has changed our lives and is essential for future prosperity.

In key stage 1 (ages 5-7), the main focus of science education is to allow students to experience and observe phenomena, taking a closer look at the world around them and encouraging them to ask questions. By key stage 2 (ages 8-11) students should begin to expand their view of the world around them by beginning to test and develop ideas more systematically. By the end of key stage 2 students should have begun to contend with more abstract ideas and make predictions about the operating world around them.

Students enter secondary school education at key stage 3 (ages 12-14) where they are more formally introduced to the three disciplines of science: biology, chemistry and physics. They should start to see the connections between the different sciences and the overarching big ideas. Students begin to work objectively, pursue scientific inquiry and construct explanations around evidence.

Upon reaching the final years of compulsory science education in key stage 4 (ages 15-16), students who wish to carry on with more advanced studies in the subject will begin to form a platform for this and establish a foundation for future career prospects. For those intending to finish their science education at the end of key stage 4, they will have acquired a level of science knowledge to understand the natural world which will stand them in good stead for an increasingly technology-enhanced society. It is perhaps in this final statement from Ofsted that expresses the importance and relevance of science education to all individuals, whether wishing to further their education or not. This echoes the sentiment expressed elegantly in the Thomson Committee report in 1918, where it was emphasised that "science specialists need to be literate and art specialists need to be numerate".

2.1.1 School Accountability

Although this study focused specifically on secondary physics education, many issues faced by physics teachers are not restricted to their department and are faced across subjects and schools in England and Wales.

School accountability is a term used to describe the quality of schools, typically in terms of performance. As with their educational structure and curriculum, accountability systems for England and Wales are different (Brill *et al.*, 2018). The most recent school accountability system in England was introduced in 2016, such measures included (Department for Education, 2018):

- Progress and attainment across eight qualifications
- Percentage of pupils entering the English Baccalaureate (EB) and average point score*
- Percentage of pupils achieving a grade 5 or above in English or Maths**
- Percentage of students staying in education or going into employment after key stage 4.¹

¹ *The EB is a combination of subjects the government deems important at GCSE. These include, English language and literature, mathematics, the sciences, geography or history and a language (Department for Education, 2017a).

^{**}Grade 5 is a "Strong Pass" in the new grading system in England and is equivalent to a B/C grade in the old A*-U system.

As of 2018, two key measures are used in Wales which include the Level 2 inclusive threshold and the capped 9 measure (Welsh Government, 2017a). The former measures the number of students who achieve five A*- C grades at GCSE including English or Welsh language and mathematics and also the Welsh Baccalaureate. The latter acknowledges nine GCSEs, five compulsory core subjects (English or Welsh language, mathematics numeracy, mathematics and two sciences) and four other subjects in which students achieved their best grades in.

Specific details on accountability measures is not divulged into in this study as it is not necessary, however what is important, is the repercussions of such systems on teaching and learning in schools in England and Wales.

The Department for Education (2018) emphasises that accountability measures are designed to encourage schools to offer a broad and balanced curriculum that rewards schools for teaching all their pupils. However, in light of other reports and enquiries into such measures, though slightly different, it is clear that in both England and Wales, the focus is entirely on performance measures. Teachers themselves have emphasised that such measures fail to contextualise schools and account for variation across cohorts. This includes factors such as areas of deprivation and special educational needs (Department for Education, 2017b). Similarly, the Wales Secretary of the National Education Union Cymru expressed that such measures lead teachers to concentrate on achieving the necessary C grades even when lower grades may represent a huge achievement for some students. Critics have echoed this statement by emphasising that such measures are incapable of recognising the hard work in some schools (Wightwick, 2018), or worse, particular low or high performing students are almost ignored as they are not a priority as they do not sit on the C grade borderline, and as a result schools are not meeting the needs of these students (OECD, 2014).

The NAHT Accountability Commission (2018) emphasise how the accountability system diverts teachers' attention from teaching and learning and directs it towards exams and performance. The Science and Technology Committee within the House of Lords emphasises that exams are largely dominated by factual recall with little testing of students' wider skills. As a consequence, less time is focused on innovative learning strategies and more on targets, progress and performance, implying that students are seen more as numbers than individuals.

Such pressures are highly influential to teachers' decision making and their pedagogical approaches that consequently encourages them to teach-to-test (NAHT, 2018; OECD, 2014). Such an approach sacrifices student autonomy and creativity and is described by Bailey (2014, p. 669) as *playing safe*. These pressures are so great and stand at the forefront of schools' accountability, that sadly education becomes largely focussed on performance, with disregard to students' learning experiences. Bailey (2014) expresses grave concern that such pressures may lead teachers to stretch from a guiding and supportive role to explicitly showing students what to

do in order to eliminate any risk of failure. The curriculum is narrowed and lessons are dominated by revision, practice exam questions and recall of facts which consequently means that anything considered to be an enrichment activity, and not directly applicable to the curriculum, is overlooked. All teaching time is used to reinforce knowledge and exam technique. One teacher in an interview with the Guardian in 2016 expressed "*The system's focus on exams means killing any creativity. Children are more than units to be tested*" (Marsh, 2016).

2.2 Applicants in Science and Physics

The Joint Council for Qualifications (JCQ) provides data for entries and results in individual subjects at GCSE and A-level across the UK constituents. Figure 1 shows the number of students who sat A-level exams in biology, chemistry and physics between 2010 and 2018 (JCQ, 2018; 2016; 2014; 2012; 2010). This data is for all of the UK but individually, England and Wales shows almost identical trends. From Figure 1 although it can be seen that applicants in all three sciences have shown a slight increase since 2016, physics remains the longstanding least popular of the three.





Entries at A-level are considered to be a valuable indicator of students' attitudes towards the individual science subjects as they can be chosen individually. Up until the age of 16 it is compulsory for students to study all three sciences to some extent. Entries in GCSE can therefore only reflect the number of entries in triple science, additional science or applied science and

therefore do not reveal much about students' preferences to biology, chemistry and physics individually. However, at A-level students are largely able to pick one, two, three or none of the sciences.

As students typically only study three or four subjects at A-level, they must choose their subjects carefully. Schools generally recommend that physics should not be taken without mathematics, it could be that this lower uptake in physics is due to the fact that students have to sacrifice two of their three or four subject choices.

Turning to applicants to University, UCAS and the Higher Education Statistics Agency (HESA) do not categorise data by individual subject and instead group them. Therefore when looking at applicants into university we can only distinguish applicants to 'physical sciences', which will consequently include chemistry subjects as well as physics.





Figure 2(a) illustrates the number of individuals in England and Wales that have applied for university since 2010 according to statistics published by UCAS (2018) where it is evident that applicants have declined in the past two years. Figure 2(b) illustrates the proportion of university applicants in England and Wales who applied for physical sciences courses. Here it is evident that the percentage of students in Wales applying to physical science courses have continued to decrease since 2014, and since 2015 in England, with the exception of a small increase in 2017.

These results indicate that not only are applicants to physical sciences in decline, but so are university applicants as a whole. This is a cause of concern for many UK businesses and the economy as is emphasised in the following sub-chapter.

2.3 The Importance of STEM and Demand for its Professions

In the UK, 8.5% of the national economic output is contributed by physics-based businesses and individually, each worker has a gross value added that is equal to twice the national average (Deloitte, 2012). However, shortages in STEM-skilled personnel remain a long-standing concern for UK businesses. Broughton (2013) estimates a skills shortage of approximately 40,000 STEM-based workers each year. The pressure on students to get good grades has led to the perception that STEM subjects are only suited to the extremely bright students.

As a result, UK employers are troubled with a growing recruitment concern and young people are at risk of not possessing the necessary skills to meet the demands of today's economy and society. In 2017 the Confederation of British Industry (CBI) and Pearson published their annual education and skills survey highlighting the longstanding problem in recruiting STEM-skilled personnel. In their 2015 report they highlighted that the number of businesses announcing difficulties in recruiting STEM graduates doubled between 2013 and 2015 (CBI/Pearson, 2015). Their 2016 report published projections of demand increases for skills across different sectors over the next 3-5 years. Engineering, science and hi-tech was forecast the highest of all sectors with an increase of 90% highly skilled personnel.

Pursuing STEM in higher education also provides excellent prospects for students as well as businesses. CBI/Pearson (2017) emphasise that their annual surveys repeatedly demonstrate that STEM graduates have an edge over the general graduate population with higher employment rates and higher median salaries.

There is clear evidence to show that the nation is presented with an issue that demands a resolution and not surprisingly, has led to a rising concern in regards to the number of qualified scientists and technologists that will be available in the near future (Osborne *et al.*, 2009). Particularly in a time where we are faced with world-wide challenges, such as climate change, we require extensive development in order for solutions to be found. Wieman (2007) emphasises that science education is no longer just important for the small fraction of the population who wish to become scientists. The global challenges that we face in the present day and age demands an entire population that is scientifically literate in order to understand, explain and defy these challenges. As it stands, we are currently presented with a vast number of individuals who are inadequately prepared for the demands of today's society and workplace (de Jong, 2006).

In 2004, the UK government responded to this ongoing problem and set out to increase A-level participation in science and mathematics by 2014. This initiative was seen to be successful, however since culmination in 2014, although uptake of mathematics has continued to rise, uptake of all three sciences has generally plateaued (JCQ, 2018). The governing bodies in education have also publically acknowledged the importance of reforming science education. Ofsted (2013, p. 2) themselves stated, "*We need better science education to secure a strong foundation for a*

successful and technological society". However, concerningly, despite this comment, the Wellcome Trust (2015) carried out a review of Ofsted reports in primary and secondary schools in England. Despite science being a core subject alongside English and Mathematics and Ofsted seemingly regarding it with high importance, Wellcome found that 93% of primary school and 33% of secondary school Ofsted reports gave no mention *at all* to science, despite 100% mentioning maths. Ofsted and Estyn are the two key educational bodies in England in Wales that govern the education system and monitor school standards. If they themselves do not place an emphasis on science in their reports then they cannot expect schools to place a high priority on their science departments.

It is clear that this problem in students pursuing science subjects and careers is ongoing and has been largely unchanged over recent years, and as demand increases, each year becomes evermore worrying. Such problems imply that it should be in the UK Government's best interest to invest in STEM education in order to encourage the growth of the economy. The following subchapter therefore explores such implications.

2.4 **Previous Science and Physics Outreach Initiatives**

According to such presumptions, in 2017 the UK Government agreed on the importance for good quality initiatives to fuel students' interest and aspirations in STEM. What's more they also acknowledged the importance of evidencing the impact of such initiatives, stating:

Having a strong evidence base that shows how effective they are is important to achieving this, and will be a key factor in teachers' decisions about what initiatives they choose to engage in and are best for their students.

The Government also claimed to ensure that government funded initiatives were appropriately monitored and evaluated. They launched the £125 million Education Endowment Foundation (EEF) to help in assessing impact and with an aim to improve education attainment and other outcomes (UK Government, 2017).

Such actions would appear extremely positive, however, review of the 'completed projects' on the EEF website revealed just three science initiatives (two secondary and one primary) compared to 42 language and literacy and 14 mathematics initiatives.

One of the secondary science initiatives, 'Let's Think Secondary Science' cost over £600,000 and aimed to develop students' scientific reasoning. The initiative was evaluated through a randomised controlled trial with over 8,000 students, but results demonstrated no evidence for improved attainment and students who were subject to the initiative performed worse in assessments than the control group (EEF, 2018a).

The second initiative, 'Spaced Learning', was a pilot aimed to boost GCSE science performance by repeating information to be learned to students multiple times with time between each repetition. Though this £300,000 project showed promise, with teachers and students appearing to enjoy and engage with the programme, (EEF, 2018b) the researcher here is sceptical. Though repetition may be effective in knowledge retainment, its revision-like procedure is not particularly innovative and the researcher is not convinced that this would promote long-term positive learning experiences among students. It is also typical of didactic teaching methods which is critiqued later on in this study.

Similar problems are recognised by the Welsh Government. In Wales, the National Science Academy (NSA) stands as the Government's primary vehicle for promoting STEM education and participation. Since October 2012 is has invested £2.2 million across 50 STEM initiatives which has delivered over 1,000 activities to over 100,000 students and CPD events to over 1,300 teachers (Welsh Government, 2015). However there are few reported outcomes and sole reliance is placed on programme output figures as a means of evaluating success.

Such evidence merely demonstrates the reach of initiatives, which although is important, tells us nothing about the actual impact such projects had on students' attitudes, aspirations or performances. The Welsh Government (2015, p.12) emphasise that an evaluation template and delivery plan is crucially needed in order to identify *"what works"*.

Nonetheless, despite the lack of evidence reported at a government level, there are some substantial efforts towards initiatives in physics, science and STEM education, with varying degrees of success. Here the researcher defines an initiative as an event that takes place in the classroom on either a short- or long-term basis and can be either internal; designed and implemented by the school and teachers themselves, or external; where practitioners, universities or businesses for example develop an initiative program and visit various schools for implementation or invite schools to come to them. This latter approach is generally referred to as outreach.

Vennix *et al.* (2018) describe outreach as a means of demonstrating the applications of classroom content to practice. Outreach should be embedded within a learning environment and should demonstrate relevance, inspire, intrigue and motivate students into the subject beyond the compulsory level (Aslam *et al.*, 2018; Turner *et al.*, 2007). Outreach can take a number of different forms including: workshops, ambassador presentations, mentoring schemes, placements, visits to industry, master classes, competitions and activities (Aslam *et al.*, 2018). Though they specifically describe these in terms of STEM outreach, they are generally applicable to any subject of outreach.

Unfortunately practitioners often report that they typically engage with the same group of schools (Aslam *et al.*, 2018), implying that often these are not the schools of the greatest concern. This is

also the experience of the researcher during her outreach work prior to this research. It was apparent from such experiences that many of the schools and teachers who were interested in outreach were not faced with so many issues, teachers were pro-active and students appeared to already display relatively positive perceptions of their learning experiences in science and physics. Aslam *et al.* (2018) also emphasise that for STEM outreach to be successful, teachers and practitioners must work together to ensure the right students are targeted and outreach activities are complementary to school activities. This recommendation will also help practitioners to understand which schools are of greater cause for concern, hosting the most disengaged students.

Nonetheless, various initiatives have been developed over previous years in attempts to improve standards in science education (Bell, 2014), at national, regional and local levels (Wynarczyk and Hale. 2009). Despite the Government's emphasis on the importance in the evaluation of initiatives, and although judgements are being made on success, Bell (2014) highlights that the majority of such initiatives, government funded or otherwise, lack the provision of high quality evidence of their effectiveness. This is echoed by Vennix *et al.* (2018) who emphasise that although outreach initiatives are widespread and demonstrate variety, knowledge and understanding of their effects is much more limited. Wynarczyk and Hale (2009) offer an explanation for this in suggesting it could be that where initiatives generally reach their targets in terms of participant reach (i.e. number of schools, teachers and students), this may be considered a sufficient indication of the success.

Poor quality evaluation also does not appear to be limited to the UK. For example, Vennix *et al.* (2018) investigated STEM outreach activities developed by companies and universities intended to inform and motivate Dutch students towards careers in STEM. The activities involved traditional lectures and workshops (spanning from one hour to week long projects), both were offered either in- or out-of-school. Their participants included 696 students aged 12 to 17 from 35 high schools in the Netherlands. Though their study demonstrated that students reacted positively to the outreach activities with positive perceptions of, inter-alia, innovation, cohesiveness, involvement and autonomy, they did not do a comparison by occasion. They therefore had no baseline for students' perceptions of the science classroom and cannot guarantee that the initiative caused positive perceptions or whether they already existed. Consequently this limits the extent to which the success of the initiative can be understood. Again although some activities span only an hour, there were multiple implementations over time, therefore again presenting the problem for teachers of finding time.

On a much more local level is the Lab in a Lorry (LIAL) program set up by the Institute of Physics in 2005 in order to address the concerns highlighted in preceding chapters. LIAL is a mobile science lab that travels to secondary schools around the UK and Ireland to deliver handson, open-ended physics experiments to students ages 11 to 14 (Grant *et al.* 2017). It is a one-off event lasting approximately three hours and therefore does not sacrifice large amounts of teaching time.

In 2017, Grant *et al.* published an evaluation report of the program. Data collection involved online surveys and telephone interviews with teachers and volunteers who delivered the outreach program, and short exit surveys with students. Such approaches demonstrated that the most in depth data was retrieved from teachers and volunteers. It was therefore immediately evident that there was not a great focus on first-hand feedback from the students themselves about their experiences. This is perhaps because more in-depth procedures would have required additional time, and therefore would have meant sacrificing students' contact time with the LIAL workshops. What is more, the volunteers would have had no prior interactions with any of the students and presumably no formal teaching experience, their observations and interpretations are therefore considered to be largely unreliable, and given their voluntary participation in the project would likely be biased to students' reactions to the initiative.

There is also research evidence that demonstrates a discrepancy between student and teacher reports of the same scenarios. That is, teachers have a tendency to report more positive perceptions about the classroom environment than their students and outside observers (Centra, 1973). Hook and Rosenshine (1979) caution that teacher questionnaire responses around classroom activities cannot guarantee an entirely accurate reflection of actual practice.

Another limitation identified among the literature was that evaluation appeared to focus largely on the specific initiative. There was little implication of the value of such findings in the long term or how such findings could be translated to other initiative designs and procedures. The little data that was collected from students subsequent to the LIAL initiative, offered little value. The short exit survey consisted of three items to which students had to respond with either yes or no, these items were:

- 1. I understood something new about the real world
- 2. I met someone with a science job before today
- 3. I would imagine doing a science job in the future

It is unclear on the purpose of question 2 and is particularly invaluable under circumstances where all students had indeed met a scientist before the initiative. It also fails to reveal any information about whether the students perceived the LIAL volunteers to be scientists and if they did, whether they saw this as an appealing career. Question 3 had no baseline for comparison, which means no inference can be made if this was an improvement from students' perceptions prior to the LIAL experience.

However, a more thorough evaluation and research procedure into LIAL was implemented by Barmby *et al.* (2008). Their methodology involved student questionnaires and focus groups.

Feedback demonstrated that students enjoyed the LIAL experience due to the practical elements and that the LIAL activities that had less of a practical focus were considered to be weaker. However another positive reported by students was that the deliverers of LIAL "*were different from our normal teachers*" (p. 1085). This point is subsequently not something that can be fed into teachers' practice and is not sustainable for reforms.

The initiative Einstein Year also carried out substantial evaluation. Launched in 2005, the Einstein Year involved a coordinated programme of outreach activities aimed at changing young people's attitudes to physics, increase quantity and quality of physics outreach and build sustainable links between physicists and their communities (Peters *et al.*, 2006).

They used a graffiti wall where students gave responses to the question 'what is physics?' before and after an outreach event. Results demonstrated increased excitement subsequent to participation, with before comments including adjectives such as '*boring*' and '*confusing*' and after comments such at '*fun*' and '*more exciting*'. Questionnaire data revealed that student participants showed a small yet consistent increase in interest in comparison to non-participants, however no significant changes were found in their opinions around the impact of science and technology on society, or around their attitudes to scientists and their work.

One particularly interesting feature of the evaluation report of Einstein Year however, is that Peters *et al.* (2006) provided recommendations for future 'year of' celebrations, therefore providing an example where results of outreach could be fed into the wider community and broader context.

Unfortunately in both cases of LIAL and Einstein Year, neither initiatives were classroom-based and both required substantial amounts of funding for implementation. Regardless of their levels of success they are not sustainable and offer limited recommendations that can be immediately or effectively employed by schools and teachers. Straw *et al.* (2012, p. 37-38) emphasise that outcomes are much more sustainable where initiatives can be implemented into schemes of work. A teacher participant in this study commented specifically:

"Big events, they look good... but they don't really live that long in the consciousness, which is why we focused on the schemes of work... it means that next year the exact same lessons can be done, and so there is some sustainability there."

More specifically, Straw *et al.* (2012, p.6) report on the evaluation of the Camden STEM initiative funded by the Wellcome Trust. The initiative involved eight schools who received £10,000 over two years to construct a STEM initiative project to address the necessary needs of their school. Though the report does not detail the specific initiatives that were developed and implemented by the school, it provides examples of successful activities and provides an

evaluation of their influence on students, teachers and the schools as a whole. Successful examples included:

- Teaching the key stage 4 curriculum in an entirely interdisciplinary way
- STEM homework exercises
- STEM enhancement and enrichment activities (including STEM clubs) on a range of themes
- The use of the school garden for experiments and data collection and analysis
- The use of pre-prepared and external OAs

Of the studies mentioned thus far, this report was considered to provide the most valuable insight into students' learning experiences. Qualitative data provided specific examples of what students perceive to be positive experiences, and initiatives were based around what schools and teachers could implement into their own practice.

Teachers reported that enhancement and enrichment activities had been successful in starting to boost students' engagement and interest in the STEM subjects. Specific comments from students detailed how the initiatives have given them a better idea of how science works in real life and that it was good to be able to apply a context to what they were learning. These comments in particular demonstrate that the initiative successfully targeted a problematic area highlighted in Chapter 2.5, yielding more positive perceptions from students as a result.

Students were also seen to react well to more interactive learning environments, reporting that they learnt more when interacting with things, getting real-world hands-on experience and having to think independently.

Though this study appears to be well evaluated and offers insight to other practitioners on how to design a STEM activity that is engaging and promotes positive learning experiences among the students, again such results were only achieved through significant amounts of funding and time provided to teachers. Given that time and money is typically reported by teachers to be among the most limiting factors to their teaching, it is considered that such positive results could be perceived by teachers and practitioners as unachievable without such funding and time availability.

There therefore remains a gap in the literature whereby although many studies have been carried out, with varying levels of success, many have simply identified where the problems lie in terms of student engagement and what needs to be targeted, but there is little recommendations of how. Other studies that shed more light are also often fuelled by large amounts of funding or demand time from teachers. Teachers need initiatives that can be directly fed into schemes of work and do not demand vast amounts of money and input from schools. Such claims are demonstrated in searches into science education initiatives that yield little evidence of substantial evaluation reports. Such problems still remain despite being voiced in previous years by Bell (2014) and Wynarczyk and Hale (2009). Where searches did obtain evaluation reports, there were lacking valuable components. For one, and as highlighted by the Welsh Government (2015), many reports merely document their reach rather than impact, others are often restricted in terms of their sources of data collection that subsequently limits the reliability of such data. Some procedures were so specific to the initiative they offered little transferable insight to other practitioners. Finally, some were indicative that in order for success, big grants are needed or someone external from the school needs to come in, thus implying reform is out of the hands of teachers.

Regardless of initiatives and their success, the problems highlighted in Chapter 2.3 are current and are only increasing in magnitude. Therefore, before identifying specifically where issues lie and what should be targeted by initiatives, it is first important to identify whether there is a particular point where problems arise in order to find a suitable target audience of initiatives. Are students' perceptions of science and physics ever positive? Or do many individuals simply have an innate dislike towards such subjects? The answers, put simply are yes and no respectively.

2.5 Identifying the Point of Disengagement

It would appear that typically, individuals do not hold an innate dislike towards science and physics. In fact, when in primary school, students present high levels of natural curiosity and positive perceptions of science (OECD, 2006; Rocard *et al.*, 2007). Instead, it is upon the transition into secondary education that students appear to become increasingly disengaged with science (Osborne and Dillon, 2008).

When entering secondary education, students are seen to possess favourable attitudes towards science, but this enjoyment and consideration of careers within a scientific field is seen to drop substantially and relatively quickly, not only in the UK (Osborne *et al.*, 2003; Barmby *et al.*, 2008; Tymms *et al.*, 2008; Galton, 2009) but on international scales (Trumper, 2006a; Rocard *et al.*, 2007; Osborne and Dillon, 2008).

Barmby *et al.* (2008) measured 932 UK students attitudes towards science (272 Year 7, 432 Year 8 and 228 Year 9). Participants were given questionnaires in which they had to respond to a series of statements with a level of agreement of 1 to 5 (1 = strongly disagree and 5 = strongly agree), where 5 indicated a positive attitude. Overall average scores were seen to be the highest among year 7s, lower in year 8 and even lower again in year 9.

More specifically, students in later stages of education demonstrated lower scores in interest in learning science in school, self-concept in science, practical work in science, science outside of school, future participation in science and importance of science.

Similarly, a study by Spall *et al.* (2004) implemented questionnaires to students in England in years 7 through to 11 in order to gain an understanding of their perceptions of physics and biology. Their results demonstrated that in year 7, students generally held positive perceptions, with 61% of students reporting they really liked or quite liked biology and 64% for physics. However, year 11 students' perceptions were substantially different, particularly for physics. At this stage 52% of students reported a liking to biology and just 29% reported a liking to physics.

Although these studies are not longitudinal, they are still indicative of a deterioration in students' attitudes towards the sciences as they progress through their secondary education. Such research is useful as it provides an indication to where things start to go wrong, but it also needs to be understood what specifically happens during secondary education that triggers this decline.

Williams *et al.* (2003) found that the three most popular reasons reported by year 10 students in England as to why they found physics boring to be due to the difficulty of the subject, a general lack of enjoyment and the subject's content, respectively. Barmby *et al.* (2008, p. 9) also followed up questionnaires with student interviews that provided insight to why students are *"switched off"*. Though their study focused on science more broadly rather than the separate strands, students reflected science is not practical, well explained or relevant.

However, though there may be a general trend of declining interest in science and physics, naturally this is not the case for all students. Although there is clear evidence that the UK is faced with problems in recruiting STEM professionals, the head count is not at zero. So what is the reason behind these exceptions? Why, when many students are being discouraged from science and physics, do some continue to persist?

2.5.1 Who Stays and Who Leaves?

Aschbacher and Roth (2009, p. 564) performed a longitudinal study in order to examine students' aspirations in science as they progress through their secondary education. More specifically they wanted to analyse the trajectories of students who were initially considered to be "*in the pipeline*" and expected to pursue science post-16. Analysis of data from interviews and surveys revealed three groups of students, high achieving persisters, low achieving persisters and lost potentials.

Almost half (n=15) of the 33 students who expressed a strong interest in pursuing science at the age of 15 were no longer interested by the age of 18, these were the lost potentials. Some influential factors identified were subject difficulty and poor school experiences. For the former point students claimed they felt they had to work much harder in science to get the equivalent grades they could obtain in other subjects with much less effort. Consequently, this pushed

students towards other subjects that they felt more capable in. This latter point was reflected in terms of poor instruction with a lack of hands-on inquiry activities and little encouragement from teachers. There was also evidence that students faced disapproval or even rejection from teachers when seeking help or guidance.

However, 18 of these 33 students continued to persist in science, 12 of these students were considered high achievers and six low achievers. Interestingly, perceptions were not entirely different and even high achievers reported that science could be hard and boring. Their continuation was seen to be due to an altruistic interest and enjoyment of understanding the world. Here, the researcher considers that such responses could imply that these students were focusing more on the larger picture and the long-term which drove their continuation. Lost potentials however were more focused on their current and immediate experiences of science.

The low achievers, although remained interested in pursuing scientific careers did slightly shift their direction and aspirations. For example one student who initially wanted to be an obstetrician and attend medical school, by the time she was 18 was looking instead at nursing programs. The researcher here considers findings of the low achievers to be particularly important. Though students were not top achievers, they had an awareness of alternative scientific avenues they could take and thus shifted their motivations accordingly.

Though the study is limited in its applicability due to a sample of American students rather than English and Welsh, a UK study carried out by Hamlyn *et al.* (2017) demonstrated similar findings, particularly in relation to the difficulty of science. Their report illustrated that students felt performance in English and Maths was more due to natural ability whereas science was more due to how hard you work. What's more, multiple studies have demonstrated the remarkable similarities in students' experiences of school science across different countries (e.g. Sjöberg and Schreiner, 2010; Lyons, 2006).

Where Ashbacher and Roth's study highlights the negative influence teachers can have, an English study by Wynarczyk and Hale (2009) emphasises the positive influence teachers can have. In an all-girls' school case study, the authors found that despite a short supply and outdated labs, many students carried science into A-level even where it was not their strongest subject. When asked, students reported this to be due to the high level of support and the approachability of the teaching department. Although the participants of this study were more relevant, it could still be criticised for its small sample and thus weakness in terms of generalisability. Nonetheless it is acknowledged for its examples of where external and vastly expensive initiatives discussed in Chapter 2.4 are not the only means of promoting science education and in fact change can be induced by subtle in-school measures and efforts from teachers. The researcher considers that what these studies lack in breadth, in terms of their sample sizes, they make up for in depth. Where many studies touch the surface of the problems faced in science education, these studies

have unpicked these problems and explored the mechanisms behind such issues. As a result they have offered much more valuable, detailed insight into the causes.

2.5.2 Summary

Chapter 2.5 has explored a narrow area of secondary science and physics education. It has revealed that individuals appear to be innately curious and interested in the world around them, yet upon reaching secondary education, the system struggles to foster the scientific interest of many students and therefore discourages future aspirations (Osborne and Dillon, 2008). Aschbacher and Roth's study offers valuable insight into students' experiences of science at school and what factors cause intentions to shift. Findings from Wynarczyk and Hale are a small but valuable contribution of evidence against the requirement for external, costly initiatives as the only source of change.

However, the researcher considers it important to explore some of the problematic areas in science and physics education in more depth. The remainder of Chapter 2 therefore explores the wider literature on secondary students' perceptions and what this means for the aims and direction of this particular study.

2.6 Secondary School Students' Perceptions of Science and Physics

Having identified that student disengagement in science and physics appears to be triggered in secondary school, in conjunction with the finding that even students who at the age of 15 held positive attitudes and intentions in science, still faced discouragement, we now turn to look at students' specific experiences of secondary school science that could be the reasons underlying such concerning trends. This chapter explores such experiences, regardless of students' initial predispositions, interests and career prospects and simply explores how science and physics is experienced in secondary education by the masses.

The literature review process began with a search of relevant literature around secondary school students' interests in science and physics both nationally and internationally, though priority was naturally given to studies based in England and Wales. What was immediately evident was the lack of research that used students in Wales as participants. Where a substantial amount of research into science and physics education in England or the UK as a whole was found, although not necessarily within the last five years, only one study was found to focus on Wales which was published by Parkinson *et al.* in 1998. Given that England and Wales have different education systems, curricula and examination bodies, it is important to gather literature on both geographies. Nonetheless, this finding; or lack of, merely emphasises the need for further research into secondary physics education in Wales. The statistics reported in Chapter 2.2 in

terms of applications into physics, post-16 level, demonstrate that recruitment in physics is just as much of a problem in Wales as it is in England.

Nonetheless, on review of the literature that was collected, much of it was seen to be reflective of the points highlighted by Aschbacher and Roth (2009) as the reason behind their "lost potential" students. As common themes were seen to emerge, the researcher began to categorise the literature into two overarching themes, evidence of positive experiences that had led to positive perceptions of physics and evidence for negative experiences which alternatively led to negative perceptions of physics. A number of sub categories were identified for each of the two strands, these are summarised in Table 1.

Table 1 – Positive and Negative Experiences in Physics Education		
Positive	Negative	
Interest and Enjoyment	Disinterested or Boring	
High Confidence	Low Confidence	
Ability	Difficult	
Relevance	Irrelevant	
Pedagogical Experiences	Classroom/School Experiences	
Usefulness	Not Useful for Future	
Career Aspirations	Don't Want to be a Scientist	
Family Influences	Family Influences	
Good Teachers	Poor encouragement from teachers	

From Table 1 it is evident that most categories appear to be almost antonyms of one another. The remainder of this chapter therefore discusses the various experiences that are reported as having negative effects and impediments on students' learning experiences in science and physics and are compared with their equivalent positive counterpart.

However, it should be clarified that many of the issues highlighted in the following sections are not standalone. That is, many of them are correlated and influential of one another which will become apparent throughout. For example, we begin by discussing students' interest and enjoyment around physics, yet many other points discussed such as confidence and relevance are intertwined with students' interest.

2.6.1 Interest and Enjoyment

Interest and enjoyment has been touched upon previously in relation to attitude, perception and experiences. Where an individual has as interest in something, the researcher assumes that they

will have had enjoyable experiences of it and hold some degree of a favourable attitude towards it. Upon review of the literature, it appeared that the most common focus of students' attitudes appeared to centre on interest and enjoyment. The Wellcome Trust (2011) found that in UK schools, one of the key elements that underpinned students' engagement in science revolved around their level of personal interest and that making science more enjoyable would subsequently improve engagement. What their report did not say however, is what makes science enjoyable.

On an international level, the ongoing 40 country collaborative ROSE project (Relevance of Science Education) gathers data around students' attitudes and motivations towards science and technology. Results from ROSE although noted slight variations across countries, identified that common across countries was students' claim that "school science has not increased my curiosity" (Sjøberg and Schreiner, 2010).

Though Sjøberg and Schreiner's study included England, Northern Ireland and Scotland among the 40 countries, the focus was on international trends, and research on a local scale implies less cause for concern. Hamlyn *et al.* (2017) reported findings on behalf of the Wellcome Trust who gathered survey data from a nationally representative sample of 4,081 students aged 14 to 18 attending state-funded schools in England. One of the headline findings appeared more promising, with 68% of students reporting that they find science at school very or fairly interesting.

However, this is not the full picture. Research into students' attitudes towards science in the UK is not clear cut. Though there are general patterns, there are a number of factors that can have a great influence on research outcomes. For one, the critical importance of distinguishing between the three core disciplines of science was identified very early on in the literature search. For example, though Hamlyn *et al.* report somewhat promising findings, when broken down into physics, chemistry and biology, the physics story is much less heartening. Spall *et al.*, (2004) emphasise that such distinctions of the sciences are not always made, with much literature referring to science as a whole. The researcher considers this to be unsurprising for two reasons. For one, science is a consistent subject throughout compulsory education, whereas it is only distinctly separated into its three core disciplines in key stage 3 in the UK or even later in some schools. It is therefore easier to gather data from a larger group of students. Secondly, it is expected that more funding is available for research in science education that encompasses a broader demand of the economy rather than the narrower disciplines.

However, physics, biology and chemistry are three different subjects and where distinctions are made, attitudinal research tells a very different story. This is immediately evident from the report from Hamlyn *et al.* (2017). Where students' attitudes initially seemed positive, when science was broken down, physics was seen to be the least popular, although this was largely influenced by

highly negative responses from girls. Chapter 2.2 also highlights that the uptake in chemistry and biology is less of a concern than physics. An attempt is therefore made to focus solely on literature that distinguishes between these separate strands of science.

Williams *et al.* (2003) explored interest in biology and physics among 317 year 10 students across six state schools in England. Where half of students reported finding biology interesting and only a quarter found biology boring, the opposite was found for their interest in physics, where half of the students found physics boring and only a quarter found it interesting.

The only research into attitudes towards science among students in Wales was published in 1998 by Parkinson *et al.* Though two decades old, results from Welsh students is largely reflective of findings from more current studies in England and other countries. Parkinson *et al.* found that attitudes to science among 1,038 13 to 14 year old students were generally positive. However, in follow-up focus groups with a representative sample of participants, students were asked about what aspects of science they liked most and between the three core disciplines, physics was cited the least among both male and female students. This finding according to gender is therefore conflictive to those of Hamyln *et al.*'s (2017).

Nonetheless, we can be confident that generally, physics is seen to acquire less interest among secondary school students than other sciences and this is particularly prominent among females.

2.6.2 Confidence versus Actual Ability

Much of the literature around students' interest in physics also highlights the difficulty of physics as an impediment to students' enjoyment. The difficulty of physics has also been reported as a strong factor influencing students' decision not to progress with the subject post-16 (Tripney *et al.*, 2010). The work of Bevins *et al.* (2005) again highlights the importance of distinguishing between the three strands of science. Where students reported that they did not find science difficult to learn, feedback from interviews revealed that physics was in fact difficult and complex; more so than biology and chemistry.

Multiple studies have demonstrated large correlations between students' confidence in their ability in a subject and their liking of said subject (e.g. Barmby and Defty, 2006; Williams *et al.* 2003). Unfortunately, common among the literature are reports from students that attribute their dislike of physics to its difficulty (Gill and Bell, 2013; Tripney *et al.*, 2010; Aschbacher and Roth, 2009).

Where students' general interest in physics is seen to decline throughout secondary education, their perception of its difficulty is seen to increase. Spall *et al.* (2004) found that in year 7, 21% reported to find physics easy whereas only 9% of year 11 students said the same. This is contradictory to biology where 20% of students claimed to find it easy in year 7 and this

percentage remained stable into year 11. This element of difficulty could imply a contributing factor as to why biology is seen to be more popular than physics.

However, interestingly, despite frequent reports regarding the difficulty of physics, in England, more students obtain A*- C grades in GCSE physics than in chemistry or biology. Physics also took the lead in Wales in 2017 (based on JCQ data, 2010-2018).

In response to such discrepancies, several researchers have subsequently looked into the relationship between perception of ability and actual ability. Barmby and Defty (2006) looked at both expected grades and actual grades in relation to the liking of biology, physics and chemistry. Though they saw a large correlation between expected grade in physics and students' liking of the subject, this correlation was not applicable for students' actual grades, and not only for physics but also for biology and chemistry.

Perhaps most worrying is the findings from Rodd *et al.* (2014) who used purposive sampling to gather data from an equal number of STEM and non-STEM undergraduates across four UK universities. All students had qualifications that would have allowed them to pursue, and be well-suited to a STEM subject at university, however they found that many of these candidates still held a perception of being "*not good enough*". Similarly, results from the Horizons in Physics Education report (HOPE) (2016) sampled gifted and talented students and even these typically reported that the difficulty of physics was the deterrent from pursuing a physics-related degree. Similar findings were found among younger students, aged 10 to 13 in the ASPIRES survey (Archer *et al.*, 2013). Interview data from 'middling achievement' students revealed that although they were highly interested in science, they felt that a career in science was "*not for them*" because they were not clever enough. Though the researcher considers that what qualifies as "gifted and talented" and the limits of what qualifications would make students "well-suited" to STEM are not specified and subjective to the researchers, such results from multiple studies are nonetheless considered to be indicative of a problem in confidence among students.

Such findings suggest it is therefore not purely a matter of ability that is resulting in a skills shortage in physics and STEM industries, the skills are held by many individuals but issues in their confidence is dissuading them from pursuing such subjects further and impeding their interest and future aspirations. Should students recognise their ability, this could help in enhancing their interest and driving continuation.

2.6.3 Relevance

Many international studies on attitudes to science education have highlighted a lack in students' perceptions of the relevance of the subject, both to themselves personally and to the world around them (Bevins *et al.*, 2005; Cleaves, 2005). In particular, OECD (2006) attribute students' declining interest throughout education to the fact that science teaching becomes increasingly disconnected from cutting-edge science and in its relevance to day-to-day life. However, given

the serious issues highlighted in Chapter 2.3, physics and science is highly relevant to our current world and so students need to be made aware of this through a more connected and applied curriculum. Such disparities have led to relevance being a common focus in reforms to science education (Stuckey *et al.*, 2013).

Nonetheless, research findings are not unanimous even among those with similar participant groups. For example Osborne and Collins (2001, p. 446) found that 16 year old students in England perceived science to be an important component of their education as it is "*all around us*" and that it helps you to "*understand the world*". Whereas Barmby *et al.* (2008) examined variations in students' attitudes in science over the first three years of secondary school in England and found that science at school is not perceived to be relevant. Though the ages of the participants in Barmby *et al.*'s study were slightly younger, if anything, their results would suggest that the participants in Osborne and Collins' study should have been even more concerning, but instead the opposite was seen.

However, there is a seven year difference between the two studies, in that time there may have been various curriculum changes if not at a national level, then perhaps on a school level in terms of structure and pedagogy. Alternatively if there had been no changes this could present an even more plausible explanation for such discrepancies. For example, if the science curriculum is largely unchanged across decades, then what was relevant and current at one point, is unlikely to hold ten years later. This consideration gains support from the report by the Wellcome Trust (2011). They explored young people's views on science education and provided a commentary from teachers. They urged that there is a great need for newer examples and applications in order to make science at school more relevant to students.

Alternatively, Reid and Skryabina (2002) reflect on how students' attitudes to physics appears to be less problematic in Scotland, they found that positive views were more pronounced in students pursuing the Standard Grade course, that was "application-led", than the Higher Grade course that has a more theoretical base. This suggests that students are more driven to continue with physics where its relevance and applications are evident than when content is abstract and theoretical. Barmby *et al.* (2008) subsequently suggest that a curriculum that is more readily geared around everyday life may be able to trigger more interest among students.

However, we are again drawn back to the familiar issue of distinguishing between the three disciplines in science as another valid explanation for contradictory results. When research breaks science down into biology, chemistry and physics, biology is typically seen to be more relevant than physics or chemistry (Osborne *et al.*, 2003; Williams *et al.*, 2003; Murray and Reiss, 2005). Such differences in studies around the relevance of science as a whole therefore cannot guarantee that participants are reflecting on like-for-like experiences.

2.6.4 Usefulness and Career Aspirations

A concept that ties closely with the relevance of physics is students' perception of its usefulness and the prospects it offers in terms of future careers. Unfortunately, much of the literature focuses on science more broadly rather than any of the three strands.

A common report among the research is that even where students hold favourable attitudes towards physics and science, this often does not translate into career aspirations (Ametller and Ryder, 2014). A report from the ASPIRES project involving over 9,000 students aged 10 to 14 in England revealed that although students generally felt that science was interesting and valuable, less than 17% would contemplate science as a career (Archer *et al.*, 2013). More disappointingly, they also found that only 43% of students who obtained an A* in GCSE physics and 30% who obtained an A continued with physics into post-16 education. These results are compatible with the previously discussed work of Rodd *et al.* (2014) and HOPE (2016) that implies even the high ability students are not continuing with physics.

However, what is remarkably apparent is how much research merely touches the surface, primarily exploring whether a career in science appeals to students and how they perceive its importance to the world. Little emphasis is attributed to students' awareness of the applicability of science and what sort of careers they perceive to involve science of any level.

The most relevant study found by the researcher was that from The Wellcome Trust (2016a). They demonstrated that students were of the opinion that science qualifications lead only to careers as scientists and had limited awareness of it applicability to non-science careers. Such findings indicate that not only does the curriculum need to highlight the application of physics to the day-to-day lives of students, but emphasis is also required around its importance to society and its application across multiple disciplines.

Osborne and Collins (2001) had begun to touch on this issue. Their focus groups revealed that students perceived science to be less valuable than English and mathematics for future employment and that science was not necessary for many careers. Although the authors highlighted that some students hinted at a recognition of the wider applicability of science beyond the traditional careers, no specific examples were yielded. However, it was noticed that one career the authors reported that students had argued science was not relevant to, was hairdressing. The researcher here recognises that hairdressing would have many scientific applications, particularly in chemistry for hair colouring.

Such an argument is therefore suggestive that research into students' aspirations in science careers is lacking in construct validity. If students do not have an informed awareness of what careers involve science then they cannot answer such questions to full capacity. Researchers not only need to be specific about their own definition of 'science careers' but they must also identify their participants' definitions and what they consider to qualify as a scientific career.

2.6.5 Classroom Experiences

A final aspect, that is perhaps more peripheral from students' subjective perceptions of physics education, is the classroom and pedagogical experiences they are exposed to. Where learning experiences are highly subjective and present difficulties in generalisability, pedagogical experiences and how science is conveyed can be much more objective and many similarities are seen from classroom to classroom (Lyons, 2006).

Classroom experiences have been touched upon as a large factor both encouraging and discouraging students from physics and science. Rocard *et al.* (2007) emphasise that although the reasons behind students' disinterest in science are complex, there is firm evidence for the connection between such attitudes and the way in which science is taught.

Trumper (2006a) and Osborne and Collins (2001) argue that the decline in interest throughout education is significantly influenced by the quality of teaching that students experience and posit that secondary school science places too much emphasis on didactic activities such as recall and copying. Although students were of similar age in both studies (15 and 16 respectively), the former study focused on participants in Israel and the latter in England, thus such results offer support to Lyons' (2006) claim of common international trends.

Similarly, Siorenta and Jimoyiannis (2008) investigated methods of physics instruction across secondary schools. They reported that physics education generally appears to place emphasis on the following:

- Teachers transmitting physics content by means of verbal explanations
- Covering the content of the specified curriculum
- Preparing students for examinations
- Strict alignment with textbooks and approved written materials
- Training students to acquire skills needed to solve conventional paper-pencil problems demanding complicated mathematic formulae.

It is expected that this didactically-dominated instruction is another negative by-product of accountability measures highlighted in Chapter 2.1.1. As emphasised by multiple authors, the accountability system encourages teaching-to-test and playing it safe. Bailey (2014) describes how it shifts teachers' approaches from guidance and support to explicit direction. A trainee teacher in Bailey's study reflected on her surprise that students expected to be told what to do rather than explore and carry out independent thinking for themselves. Such pedagogical approaches in physics reported by Sorienta and Jimoyiannis are demonstrative of this teach-to-test approach.

These points are reinforced by student focus group feedback in the report from Osborne and Collins (2001). Sixteen year old students in England reflected that lessons are largely dominated
by processes of copying notes from the board or from textbooks and opportunities for practical activities were rare. This method was also described to be both boring and not useful to their learning. One student in the study also commented on the difference in the nature of practical lessons and note copying lessons. Where students were able to participate in practicals it triggered them to ask questions and hold discussions that generated interactivity. However where students have to copy notes, they are so rushed to write everything down that they get no real sense of what they are writing and no real opportunities to ask questions.

Similarly, on an international level, Lyons (2006, p. 591) reported specific comments from students explaining how science was often conveyed to them in lessons:

"This is it, this is how it is, and this is what you learn."

"It is like this, learn it because it is right, there is nothing to discuss."

"It happened, accept it, you don't need to know this until A-level."

Two large scale studies ran by the Wellcome Trust both identified 'having a bad teacher' as the most commonly cited factor for discouragement in science (Tripney *et al.*, 2010; Hamlyn *et al.*, 2017). However, the same two studies and that of Mujtaba and Reiss (2013) also highlight the positive influence teachers can have on students' attitudes and intentions to participate. Such reports are indicative of the substantial influence teachers have over students' learning experiences both for the better and worse.

However as with students' future aspirations, some research into classroom experiences appeared to be limited to a surface level. Though they report the influence of having a 'good teacher' or a 'bad teacher', they do not always highlight specifically what constitutes a 'good' or 'bad' teacher. This is not to say that such definitions do not exist, Coe *et al.* (2014, p. 2) define great teaching as *"that which leads to improved student achievement using outcomes that matter to their future success"*. At face value, this definition appears to be relatively holistic, yet Coe *et al.* accompany their definition with six measures of great teaching which breaks this definition down into individual components:

- 1. Pedagogical content knowledge
- 2. Quality of instruction
- 3. Classroom climate
- 4. Classroom management
- 5. Teacher beliefs
- 6. Professional behaviours

Several studies are seen to focus on one or some of these components, for example Fraser and Fisher (e.g. 1982) have focused a lot of their research on classroom climate in Australian classrooms. Unfortunately no UK study was found to explore the complete definition of great

teaching provided by Coe *et al.* at secondary school level. Goodrum *et al.* (2001) composed a Secondary School Science Questionnaire for students which included a dimension on teachers' actions. This dimension or questionnaire as a whole has been used in several studies (e.g. Bartlett *et al.* 2018; Danaia *et al.*, 2012), though these are primarily in Australia.

A UK study by The Wellcome Trust (2011) however sheds light on what students perceive to be positive teaching. Students reported that they did not enjoy science under circumstances where the teacher did not make learning fun or when there was a lack of discipline in the class. Alternatively, where teachers were good at explaining things, students found science more enjoyable.

Osborne and Collins (2001) also retrieved more in depth data via their focus groups with students. They identified a general consensus among students that interest was sustained by teachers who made lessons 'fun' through their methods of presentation and through carrying out practical activities. Students also valued attention from teachers and those who were happy to explain things and help students when they faced struggles. This comment was particularly common among students who claimed to have teachers that provided opportunities for their students to have an active role in their own learning.

As well as highlighting the 'good' teachers, students also reflected on qualities of a 'bad' or 'weaker' teacher. These teachers typically resorted to writing on the board and working from textbooks and did not explain things fully when students did not understand them which they claimed offered a "*better opportunity for learning*" (p.460). Such findings from Osborne and Collins' study, which has been reflected on several times, are again reflective of the teach-to-test approaches resulting from accountability measures and offer valuable insight into what is happening in classrooms in England, reporting directly from the students themselves. Where many quantitative-based studies offer a general grasp on the big picture and where the problems lie, due to a lack of depth, they offer little in terms of evidenced recommendations for improvement. Although qualitative studies typically involve a much smaller sample, what they lack in numbers they arguably make up for in the depth of knowledge and understanding they reveal.

Teacher influence is also not limited to their academic impact, evidence is also demonstrative of the social influence teachers can have. Pianta *et al.* (2012) highlight that student-teacher relationships are a reflection of a classroom's capacity for encouraging development and a key to understanding student engagement. Where positive relationships can enhance various aspects including students' attention, applied effort and affective processes, negative or hostile relationships can inhibit them. More specifically, Roorda *et al.* (2011) highlight the various research that has demonstrated significant associations of student-teacher relationships and the social functioning of students, their behaviour, engagement and academic achievement.

Such influences were evident in the girls' school case study from Wynarczyk and Hale (2009) detailed in section 2.4.1, where students cited supportive and approachable teachers as their reason for pursuit. Although this was just one school, it nonetheless provides evidence of the value of student-teacher relationships that came directly from the students themselves. Or alternatively, Mujtaba and Reiss (2013) worked on a much larger scale, with 5,034 year 10 students in England. Similarly, they also found strong correlations between students' intentions to participate in physics post-16 and whether students believed that their teachers wanted them to study physics and were good at explaining physics.

Such evidence is therefore indicative not only of the importance of teachers' pedagogical approaches, but also of the relationships they form with their students. The researcher therefore acknowledges the importance of considering both aspects within a classroom and both are subsequently discussed in greater depth in Chapters 3 and 4.

2.6.6 Summary

Accountability measures in England and Wales have been widely criticised for their primary focus on performance and student achievement of A*- C grades. As a result, teachers are often forced to teach-to-test and play safe in their approaches by applying didactic pedagogies in the classroom.

Various initiatives have been implemented with an aim of promoting students' attitudes or achievement in science but many efforts are confined by their lack of robust evaluation, repeat schools that aren't necessarily the ones that need targeting, high costs and their take up of teaching time for often what could be considered enrichment activities rather than curriculum enhancement.

Such initiatives present the following gaps:

- A need for a robust evaluation framework for science education initiatives.
- Outreach activities that are parallel to the compulsory curriculum and therefore appeal to a wide audience of teachers.
- Outreach activities of which success is not conditional on additional funding, long-term interventions or visits from individuals external to the school.

Chapter 2.5 identified that individuals do not have an innate disinterest in science but that their attitudes decline throughout secondary school. This therefore presented a suitable target audience for this study.

Finally, Chapter 2.6 though not exhaustive, explored some of the key underpinning influences on students' perceptions of their learning experiences in physics and discussed research both at a national and international level. It is clear that whether a study focuses on a single classroom in

England or thousands of students across multiple countries, similarities are reported in results and the most profound issues are clearly evident.

Review of the literature demonstrated the importance of distinguishing between the three separate strands of science and how findings can often be misleading when this is not accounted for. Although many studies have often demonstrated favourable perceptions of 'science', a very different picture is seen for the three core disciplines. Where favourable attitudes are typically focused around areas of biology which is considered more relevant and comprehendible, physics is deemed by many secondary school students as uninteresting, irrelevant and difficult. Such findings therefore presented a focus for this study and specifically what areas need to be targeted in order to attempt to promote positive learning experiences among students.

2.7 Aims of the Study

This chapter has highlighted the importance of and problems with physics (and more broadly, science) education in England and Wales, and some of the initiatives that have been implemented in response. Review of the literature has revealed that in terms of physics, in general terms, we are no further forward and poor attitudes and a lack of uptake of post-16 education still remains. Several gaps were also identified in the literature around previous interventions in physics, science and STEM education which pave the way for this study.

As heartening as it is to see that attention is being attributed to current problems, many efforts are suppressed by their lack of robust evaluation or offer of transferable recommendations that can be fed into the broader context and can induce changes to physics teaching practices and enhance schemes of work. Often such efforts provide little insight to practitioners seeking to develop new initiatives, or to teachers who did not participate in the particular intervention. Alternatively, where research does report valuable evidence, their impact is often limited due to their substantial funding which, generally, is not readily available to schools and may consequently lead teachers to believe that such initiatives or improvements cannot be successfully implemented without it.

Accountability measures mean that teachers are under extreme performance pressures and any activity that is considered merely enrichment, without explicit application to the curriculum, is not compatible with such pressures. Practitioners must acknowledge such affairs if they are to be successful in addressing both some of the negative learning experiences among students and in producing something that is perceived to be worthwhile from a teacher's perspective and potentially feed into their schemes of work. This study therefore set out to identify and understand what constitutes an 'effective outreach activity'. In this context, the researcher defines an outreach activity (OA) as an activity that is developed externally to the school and teachers, but can be implemented by teachers into their set scheme of work. Understanding what constitutes an effective OA is achieved through detailed evaluation of a series of pre-existing

OAs in terms of their influence on students' learning experiences in physics, teachers' implementation, and how they tie in with current issues in physics and education as a whole faced by English and Welsh schools. It is emphasised that this study did not set out to reinvent the wheel and solely identify whether such specific OAs are effective. Instead of saying this OA caused x, y and z, this study delves into the broader context and takes the approach of saying aspects a, b, and c of a given OA are capable of causing x, y and z under w conditions, this is indicated by students and teachers as the primary stakeholders. This aim is what guided the specific methodology of the study that is set out in Chapter 6. Where a study seeks to understand the interaction between a, b, c and x, y, z, an objective positivistic approach is not suitable. Instead the research is constructivist and the researcher becomes very much a part of the research process and environment.

This study investigates specifically what features of an OA stimulates positive or negative learning experiences, and consequently what this implies to practitioners when developing classroom initiatives and to teachers when employing them or designing them themselves. More specifically, the research questions addressed in this study are fourfold:

- 1. How are students' perceptions of positive learning experiences promoted within a physics lesson?
- 2. How does teachers' planning and delivery of the Down2Earth and Hubble outreach activities influence students' learning experiences?
- 3. How does the implementation of the Down2Earth and Hubble outreach activities reveal the opportunities and limitations they lend to students' learning experiences?
- 4. What are the implications for the future development and utilisation of physics outreach activities for developers and teachers?

Given that focus is attributed to students' learning experiences and it is their perceptions towards physics that are the root of the problems faced, data collection is primarily targeted at students. Student feedback and behaviour stands as the main source of evidence around students' learning experiences. A robust, detailed evaluation procedure is constructed that is not only used to assess the impact of this study on students' learning experiences but can also be used to guide evaluation procedures of other educational initiatives in science education. Having a detailed understanding of exactly what aspects and activities engage students and stimulate their curiosity, provides invaluable insight into the future direction of science education and could help to overcome barriers that we currently face in encouraging students to pursuing physics and science in the future. A detailed description of the design and methodology of this research is set out in Chapter 6.

Chapter 3

Instructional Design

Instructional design theories are used for guidance and support on how individuals learn and develop their education. They can encompass various processes such as *cognitive*, *emotional*, *social*, *physical* and *spiritual* (Reigeluth, 1999, p. 5).

Ertmer and Newby (2013, p. 44) specify a series of questions that practitioners must consider when developing an instructional method. These are fourfold:

- What are the situational and contextual constraints of the application?
- What is the degree of individual differences among the learners?
- What form of solutions will or will not be accepted by the learners as well as by those actually teaching the materials?
- What are the learning problems?

The authors emphasise that a practitioner cannot recommend an approach or reform to instruction without a clear diagnosis and understanding of the fundamental problem and where current instructional designs are falling short. That is, the researcher here must first recognise the problems faced in physics instruction in England and Wales before she can attempt to provide an OA that effectively addresses such problems and targets appropriates aspects of students' learning, and develop a framework for evaluating its effectiveness.

The previous chapter discussed the various obstacles faced by physics teachers, and touched upon some of the classroom experiences and pedagogies employed. This revealed that the teaching of physics content appears to be largely transmissive and didactic in nature.

In choosing an instructional design method for the basis of the OAs, the researcher had to identify the current approaches typically employed in physics instruction and how and why they are hindering students in the above areas. Once this is achieved, the researcher can then consider the method of instructional design that best accommodates for such problems.

3.1 Learning Theories

There are three primary learning theories that have evolved with advances in educational research: behaviourism, cognitivism and constructivism (Ertmer and Newby, 2013).

Central to behaviourism is the learner's response to environmental conditions and cues. Learning occurs in the form of a response to stimuli and future recurrence is encouraged through positive reinforcements (Cooper, 1993; Ertmer and Newby, 2013). Learning is therefore considered to be directly observable and takes place through classical conditioning as proposed by Pavlov in 1902.

However, as practitioners and psychologists began to focus on more cognitive processes, several confinements of behaviourism were recognised and consequently triggered the development of cognitivism. Although cognitivism still places an emphasis on environmental conditions, the focus is shifted from *what* students know to *how* they know it. Where behaviourism jumps straight from the stimulus to the response, cognitivism is concerned with the process between these two points in terms of how information is received, organised, stored and retrieved in the mind (Ertmer and Newby, 2013).

Today however, many educational researchers argue that although important, less focus should be attributed to the learner's environment and more attributed to the learner themselves. Where behaviourism and cognitivism consider learning to be objective; where the world is separate and external to the learner, this modern theory regards knowledge to be subjective; a function of how an individual interprets their environment and creates meaning from such exposure and experiences. This theory is called constructivism. Where learning theory moves from behaviourism to cognitivism to constructivism, the focus of instruction shifts its emphasis from teaching to learning (Ertmer and Newby, 2013).

Constructivists place an emphasis on context and consider learning to be situationally determined. Transfer of knowledge will not take place unless learning experiences are embedded within meaningful contexts which the learner can explore and interpret for themselves. However, Chapter 2.6.5 highlighted that many physics teachers, to a considerable extent, still apply more traditional approaches that are largely reflective of behaviourism. A search of literature around constructivist approaches from teachers yielded little research, this echoes the claims of Mulhall and Gunstone made in 2012. Such observations are considered by Tsai (2002) to be due to the minority of teachers actively employing such views. That is, as emphasised in Chapter 2, classrooms are often characterised by teacher-centred, transmissive approaches rather than student-centred, experiential approaches. As a result, students are not offered substantial experiences to construct their learning for themselves.

Though attention in the field of educational research has been dominated by constructivism only more recently, its foundations were laid in the work of John Dewey, Lev Vygotsky and Jean Piaget in the early 1900s. Clements and Battista (1990) although focus on constructivism in mathematics education, provide relevant arguments to science education. They posit that constructivism allows students to develop learning processes that are more complex and powerful than those that result from traditional teaching methods, thus increasing their capability of problem solving. In addition, where transmissive, behavioural approaches present students with pre-defined knowledge, constructivism encourages students to seek such knowledge and employ their own methods and processes to explore it. This allows students to become autonomous and self-motivated in their learning (Clements and Battista, 1990), characteristics that DeVries (2002) considers to be impossible under transmissive relationships between students and teachers. The research from DeVries (2002) demonstrates that where students are exposed to sound implementation of constructivist programs, they perform better academically, personally, morally and socially.

In light of the definitions of constructivism throughout the literature, the researcher appointed her own combined definition of constructivism that is followed in this study. Here constructivism is recognised as a process whereby learning occurs as a result of the learner's own contextual experiences. Learners are not presented with a number of facts that they must take as bible, but instead must experience the processes that yield such facts in order to acquire knowledge and understanding. The focus is shifted from statements of knowledge, which is pre-specified to knowledge acquisition that cannot be pre-specified (Karagiorgi and Symeou, 2005).

However, having praised the benefits of constructivism, given that teachers often appear to be resistant to constructivist approaches, it cannot be without its disadvantages. For one, as individuals' experiences are subjective, a common set of outcomes for every learner in the classroom cannot be guaranteed. Similarly, it does not account for students' initial understanding or any misconceptions they hold (Karagiorgi and Symeou, 2005). The researcher acknowledges that this could be particularly problematic to teachers given the issues of accountability measures discussed in Chapter 2.1.1. Where schools are pressured by performance, the teacher would presumably approach a lesson with a set of learning objectives that they would aim for all students to obtain by the end of the lesson. Any lesson that risks students obtaining different understandings which may not be exactly parallel to their exam, would not be preferable.

The shift of control from the teacher to the student also raises concerns. Karagiorgi and Symeou (2005) highlight that too much control can risk the assumption that students will learn or will acquire the correct understanding, that is, they could construct an incorrect understanding from their experiences.

Edelson *et al.* (1999) also insist that students need to be motivated towards the topic of exploration in order for meaningful learning experiences to take place. The researcher considers that such motivation is arguably even more vital in constructivist learning environments than behaviourist ones as students are given more control and need to take more of an initiative. With more responsibility over their own learning, students need to be curious and want to understand the subject matter.

However constructivism is an umbrella term, and can be applied through multiple approaches. The following chapter discusses one particular form of constructivism that the researcher considers to be proficient in all the advantages of constructivism but also capable of curbing some of the limitations that have been raised.

3.2 Inquiry-based Learning

One particular teaching approach rooted within constructivism and has gained momentum in recent years is inquiry-based learning (IBL). The inquiry-based approach is reflective of constructivism in that it is based upon the awareness that scientific concepts are only fully understood under circumstances whereby the learner has constructed their own thinking and experiences (InterAcademies Panel, 2010), and is responsible for the acquisition of their new knowledge (Pedaste *et al.*, 2015). Learning is achieved through inquiry of the student and focusses on a search for an understanding and explanation of a concept in order to satisfy a curiosity (Haury, 1993), rather than through the presentation of scientific content from the teacher or a textbook (Levy *et al.*, 2011a).

The inquiry process stimulates the curiosity of students and actively engages them with the challenges and problems that are presented by the research scenario (Levy *et al.*, 2011a). By creating an emotional response that triggers interest and attention, IBL proves both more relevant and exciting than learning through deductive means and is capable of promoting a much more positive attitude towards science (InterAcademies Panel, 2010).

This approach has been praised not only by educational researchers but also governing bodies in the UK. In 2014, the Department for Education (p. 3) in England stated that science should be taught in a way that gives students an "*appreciation of the relevance of science in their everyday lives*" through "*working scientifically*". Two particular points they raise is that secondary school students should:

"Develop their understanding of the nature, processes and methods of science, through different types of scientific enquiry that help them to answer scientific questions about the world around them."

"Develop and learn to apply observational, practical, modelling, enquiry, problemsolving skills and mathematical skills, both in the laboratory, in the field and in other environments."

Similarly in Wales, students are expected to understand the role of science and scientists in society (Welsh Government, 2008). They should be capable of working both individually and in groups to investigate relationships between data, evidence, theories and explanations and develop

their skills in practical work, problem-solving and inquiry. Estyn (2017) also report the positive effects of such opportunities and the benefits of group and class discussions to students' learning.

Between 2010 and 2013, Ofsted inspectors (2013) visited 91 primary schools and 89 secondary schools, including 53 sixth form colleges and six special schools in order to assess the standard of science education. They reported that schools who reflected the best teaching of science set out to encourage students' natural curiosity and held scientific inquiry as the foundation. Problems appeared evident among schools who were too attentive to exam results and did not seek to establish student understanding or aim to apply scientific ideas via practical and inquiry-based methods. This latter point is considered to be parallel to pressures of accountability. Similarly in an earlier report (2011), they emphasise that the highest standards were seen in schools where there was evidence for students' regular involvement in all aspects of science investigations, such as forming a hypothesis, planning and performing an investigation and then evaluating their findings.

Estyn's most recent thematic report in science explored standards at key stage 3 and 4 (2017). One of their key recommendations was for schools to provide both stimulating and challenging learning opportunities with practical work involvement for students of all abilities.

As with constructivism, the concept of IBL is not a new one, and originated from the work of three theorists several decades ago, these were David Ausubel (1963), Jean Piaget (1977) and Lev Vygotsky (1978) (Cakir, 2008; Minner *et al.* 2010).

Loucks-Horsley and Olson (2000) also highlight Schwab's (1960) large influence on the initial foundations of inquiry-based learning. He proposed that when approaching a topic, teachers should employ real laboratory experiences to lead their educating, rather than following day-to-day classroom procedures. This is echoed by Wood *et al.* (1976) who explains that teachers should approach their lesson in the same way they would approach an experiment, with a clear goal and methodology. This approach allows students to first experience the science before they are presented with the underlying concepts. As a result, the final explanations proposed by students are built upon evidence from their experiences rather than information they have merely been given.

3.2.1 The Case for Inquiry-based Learning

Having set out the foundations and theories behind inquiry-based learning, we now turn to the characteristics of an inquiry-based classroom, its opportunities and limitations. Such characteristics not only set the basis on which the OAs were developed around but will also influence the evaluation protocol in order to identify the extent to which the OAs were successful in implementing elements of inquiry into the classroom. Spronken-Smith and Walker (2010, p. 726) describe the following key characteristics of inquiry-based learning:

- Learning is stimulated by inquiry that is steered by scientifically oriented questions or problems
- Learning is based on a process whereby knowledge is constructed through new understanding
- Learners are 'active' in the process and learn by doing and experiencing
- Teaching becomes student-centred and the teacher acts as a facilitator
- Learning is self-directed by students and they have increased responsibility and control over their learning.

The key characteristic of inquiry-based learning that is considered to aid in reducing the limitations of constructivism highlighted in the previous chapter is that it can encompass a variety of levels. Such levels are typically dependent on the amount of control that is assigned to the students (Spronken-Smith and Walker, 2010). Thus, before constructing the OAs prior to this study, the researcher had to decide which level of inquiry would be most suited to the audience of this study and was also able to accommodate for some of the concerns associated with constructivism.

If a continuum were to place traditional instruction at one end, signifying an entirely teachercentred classroom where students acquire knowledge through transmission of information, on the opposite end we have complete constructivism. Here the lesson is entirely student-centred, information is constructed entirely by the learner (Kirschner *et al.*, 2006), inquiry is open ended and students must construct the question or problem themselves. This level therefore demands the highest cognitive ability from students (Banchi and Bell, 2008). There are then several levels of inquiry that are commonly recognised and situated between these two extreme pedagogies. Each level of inquiry varies from the next in terms of how much control and responsibility is assigned to the students and how much guidance and support is provided by the teacher. Such levels are commonly referred to as confirmation, structured and guided inquiry. (Banchi and Bell, 2008; Bell *et al.*, 2005; Herron, 1971; Schwab, 1962; Tafoya *et al.*, 1980; Protopsaltis and Specht, 2013). The fourth level is open inquiry which describes the extreme end of the inquiry spectrum.

Confirmation is the lowest level of inquiry and sees the teacher guide students through an entire activity (Protopsaltis and Specht, 2013). Students are given a research question, its answer and a procedure to follow in order to prove and explain this answer (Banchi and Bell, 2008). Confirmation introduces students to scientific inquiry, enabling them to practice the basic skills that are involved, largely making observations and recording and analysing sets of data.

Under structured-inquiry, students are still provided with the research question and procedure, but it is now up to them to construct explanations that complement the evidence they collect through following the method they are given (Banchi and Bell, 2008). For guided-inquiry, students are presented only with a research question; it is now their duty to design an experiment and based on their findings, construct appropriate conclusions to this question (Spronken-Smith and Walker, 2010). A summary of the various levels of inquiry and how much guidance is provided to students is provided in Table 2 and adapted from Banchi and Bell (2008).

Table 2 – Variations in Levels of Inquiry adapted from Banchi and Bell (2008, p. 7).				
Level of Inquiry	Provided to Student			
	Question	Procedure	Answer	
Confirmation	Х	Х	Х	
Structured	Х	Х		
Guided	Х			
Open				

3.2.2 Optimum Level of Inquiry-based Learning

However, although researchers are largely in agreement with the advantages of inquiry over more traditional methods of teaching, disagreements are evident among which level of inquiry is most advantageous and how much structure and guidance students should receive in order to achieve the most beneficial learning experiences.

Many issues were also identified across the literature. For one, a lack of consistency in the definitions of each of the inquiry levels was clearly evident. Although the specific definitions applied in this study have been set out here and are reflective of those of Banchi and Bell (2008), other researchers attribute alternative definitions.

For example, Sadeh and Zion (2012) investigated attitudes towards guided and open inquiry projects in biology among 295 16-19 year old students in Israel. Results from an attitudinal questionnaire revealed that students who completed the open inquiry project reported higher levels of satisfaction and gained greater benefits than those who completed the guided inquiry project. They also claimed to have felt a greater sense of involvement in the project and more cooperation with peers.

Chatterjee *et al.* (2009) however did a similar investigation on undergraduate students in Texas enrolled on a chemistry course. They contrarily found that 77% of their participants claimed they would choose a guided-inquiry lab over an open-inquiry lab, 46% felt they learnt more and 43% felt they got better grades. Such results are illustrated in Figure 3. Remaining percentages are participants who were neutral in their responses.

However, at a closer look, where Sadeh and Zion's definition of guided inquiry is parallel to Banchi and Bell's and this study, Chaterjee *et al.*'s definition is what Banchi and Bell describe as structured inquiry. Such disparities were evident across multiple studies in the literature. Another issue, though more specific to this particular study was that many studies were seen to focus on biology and chemistry instead of physics in terms of subject and undergraduate students instead of secondary school students in terms of audience.

Also of interest is that despite the push from Ofsted and Estyn, little research into inquiry-based learning in science classrooms in England or Wales was found in the literature search.

In response to such issues, much of the decisions for what level of inquiry deemed as most suitable to this study was largely based on inferences made by the researcher and in light of pressures on schools and teachers in the UK.

However, irrelevant of definition, age group and subject is the study by Zion and Mendelovici (2012). They suggest that inquiry should be a progressive process and that students should move gradually from more structured processes towards open inquiry. When turning to literature that explores the effectiveness of open-inquiry when placed as a direct jump from didactic teaching, similar arguments are seen, and research evidence is suggestive of the improvement of inquiry abilities over time and experience. Spronken-Smith and Walker (2010) for example although argue open-inquiry as the most advantageous level, also emphasise the importance of more structured levels in equipping students and developing their skills ready for more advanced, open levels. They assert that implementation of the lower levels into the science curriculum should not be undermined by the positive results from open-inquiry.



Figure 3 – Students Preferences of Inquiry. Summary of Results from Chatterjee et al. (2009)

Prince and Felder (2006) parallel this progressive technique emphasising that jumping straight into open inquiry will not automatically lead to increased satisfaction and performance among students. In response, the researcher therefore expects that by undergraduate level, students would likely have more experience with scientific inquiry and thus feel more comfortable and capable in open inquiry. Younger, secondary school student however with less experience would benefit from additional structure. Arslan (2014) concludes that open-inquiry is only beneficial where students have prior knowledge of the process of inquiry and have progressed their skills. In the absence of this, open inquiry is simply too difficult and must develop from more supportive, guided levels of inquiry.

3.3 Limitations of Inquiry-based Learning

But again, as with constructivism, despite evidence for success, many studies still report that teachers typically apply more didactic, transmissive approaches in their classroom (Osborne and Collins, 2001; Lyons, 2006; Siorenta and Jimoyiannis, 2008). Even reports from Ofsted (2013) and Estyn (2017) have emphasised the importance of inquiry based learning and that not all schools are giving it priority. Particularly in Wales, Estyn emphasised the concern that practical investigations are increasingly seen to be replaced by worksheets and videos, thus restricting opportunities for students to develop essential scientific skills.

The researcher therefore reviews the potential shortfalls of IBL before specifying the particular approach that is adopted in this study in light of the advantages and disadvantages discussed.

For one, the issues in accountability measures discussed in Chapter 2.1.1 highlight why performance pressures are encouraging teachers to teach-to-test rather than stimulating positive learning experiences, which consequently does not marry with the nature of IBL. Songer *et al.* (2003) highlights that didactic techniques are easier to implement and of lower risk than more indepth and involved lessons involving inquiry, investigation and practicals. The pressure of performance standards consequently often overshadows the higher-risk approaches to teaching.

Another problem faced by schools is recruiting physics teachers. Chapter 2.1.1 highlighted the more general shortage of teaching staff and the Department for Education (2017b) mentioned physics as a particular subject that faced this problem. More specifically, in 2015/16 just 71% of positions on physics teacher training courses were filled (National Audit Office, 2016), thus demonstrating that there are not enough physics teachers to meet the requirements of today's schools. Matters had worsened by 2017 where such courses saw a 19% drop from 2016 (Ward, 2018).

As a result of such deficits, schools have suffered and students are consequently taught by teachers who are not physics specialists. In 2015 the Department for Education revealed that just 67% of physics teachers in schools held a physics degree and therefore approximately a third of physics teachers held no post A-level qualification in physics.

In an interview with TES (formerly the Times Education Supplement), the president of the Royal Society of Chemistry, Sir John Holman claimed this to be a reason for the abundance of didactic

teaching and lack of practical science lessons. He asserted that specialist teachers have a greater sense of confidence and preparation in implementing experimental lessons (George, 2017).

Such a claim is reflective of reports from the InterAcademies Panel (2010) who posit that the pedagogical approach employed by a teacher is highly influenced by their subject knowledge. The fear that discussions of particular topics might extend beyond their area of expertise and comfort zone deters them from more open-ended methods such as inquiry. Teachers instead take the 'safest' pedagogical approach, which typically rely on textbooks and do not engage students in discussion or inquiry of the subject. This safest approach is also encouraged by accountability measures highlighted in Chapter 2.1.1.

The benefit of specialist teachers is demonstrated by the case study reported by Wynarczyk and Hale (2009). This school saw all three disciplines of science within the top five most popular subjects at A-level and also employed teachers who taught almost exclusively within their specialist subject. Though this does not demonstrate causation, a relationship is evident.

A third potential barrier to implementing IBL is time constraints. It is recognised that inquirybased methods can be a lot more time consuming and present more work and planning for the teacher than traditional approaches. Gonzales *et al.* (2004) reports findings from the Trends in International Mathematics and Science Study (TIMSS) that found teachers to be deficient in the time they have available to cover the entire curriculum content with sufficient depth. As a result they select the easiest and most familiar method of teaching. Similarly a study by the Wellcome Trust (2016b) reported a lack of time to be a common barrier reported by teachers. The researcher considers that these findings could suggest that even under circumstances where teachers may wish to implement more IBL, it is simply not feasible under the strict time constraints.

3.4 The Inquiry Approach for this Study

The researcher's motivation in structuring the OAs that are evaluated in this study resulted from the literature around the different levels of inquiry and also the barriers to reform that are faced by teachers.

Complete open inquiry was ruled out for several reasons. For one, it is generally recognised that secondary students cannot be expected to design and carry out a full experiment from start to end (National Research Council, 1996; Loucks-Horsley and Olson, 2000; Banchi and Bell, 2008). A consistency is seen across ages from primary to post-secondary education (Simons and Ertmer, 2005) in that the skills of scientific inquiry are not instinctive for many students (Hung, 2011). Alternatively, the key to successful scientific inquiry is attributed to the guidance that students receive in the early stages of being introduced to inquiry (Simons and Ertmer, 2005) and on the

condition that they have previously undergone more structured practices with appropriate scaffolding (Hung, 2011). These experiences will allow students to develop their capabilities in inquiry (Roblyer *et al.*, 1997) and progress towards the more open levels.

Another influential factor to the researcher's decision ties back to students' confidence. It is assumed that students will be less likely to enjoy and engage in an activity that they identify to extend beyond their capability (Fitzgerald, 2015a). This is supported by the work from Prince and Felder (2006) who found that if students are given increased responsibility over their learning, they will quickly become resistant if they are not provided with adequate support. This resistance is accompanied with a potential to turn into hostility which if experienced by teachers, will deter them from continuing to apply this pedagogical approach. Simons and Ertmer (2005) suggest appropriate scaffolds and guidance should be incorporated within inquiry and adjusted in accordance to the level of difficulty that is suitable and achievable to that student.

The OAs were therefore designed to align with, though not be limited to, structured inquiry, providing students with the research problem and the process of investigation (see Appendix 1 and 2 for OA extracts). This level was chosen as the researcher considered it to be most suited to the target audience of the OAs and most proficient in overcoming the disadvantages of constructivism and IBL. However, the researcher did explain to teachers that the level of inquiry they implemented was at their discretion. This decision is supported by the work of Levy *et al.* (2011b) who emphasise the importance for flexibility in this pedagogical approach and that teachers should tailor their level of inquiry to the context of their classroom and the needs of their students. The researcher considered that structured inquiry was also most suitable in this respect as it would be easier for the teacher, particularly if they were not a physics specialist, to have all the structure and support already embedded in the OAs and merely had to remove aspects in order to tailor the activity to their students rather than add things in and thus requiring additional work from them.

Table 3 summarises the disadvantages of constructivism that have been expressed and how structured inquiry and the approach of the outreach activities in this study overcome or at least limit such problems.

Limitation	Method to Overcome
Risk of different outcomes for different students	A structured level of inquiry provides students with the research question and the procedure but not the answer. Given this substantial level of guidance and a question to work around, it is less likely that students will yield largely different findings or understandings.
Lack of accountability of students' initial understanding	The OAs are designed so that students are presented with a research problem which they must first discuss and consider related concepts in order to construct a hypothesis. The teacher can therefore use this initial activity to focus and guide students' preconceptions and expectations.
Students must be motivated	This element is not directly related to the inquiry process itself and is instead addressed through contextualisation. This is described in the following chapter.
Time constraints	Though teachers are faced with extreme time constraints and struggle to cover all curriculum content, the approach for this study is structured enough for students so that it doesn't require as much time as open inquiry, and OAs are also directly applicable to curriculum content so should fit into schemes of work and thus not require any additional lesson time.
Lack of subject specialists in physics	Structured inquiry confines the number of avenues that students can take in their exploration and discovery. Teachers will therefore have a clearer idea of the possible outcomes of such processes and should not be daunted by unexpected outcomes.

Table 3 - How Structured Inquiry Tackles Barriers to Constructivism

3.5 Summary

This chapter has set out the various learning theories and approaches to instructional design. Advantages and limitations of constructivism as a whole, as well as IBL have been reviewed in conjunction with consideration of problems faced by schools and teachers. This process led to the decision for OAs to employ a structured level of inquiry. In line with this decision, the OAs were designed to present students with a research question and procedure, but it was down to them to complete the procedure and draw conclusions based on their experiences.

Having set out the instructional design of the OAs and justified such structure, this chapter has primarily addressed some of the issues seen in classroom experiences and to some extent, in students' confidence. The following chapter divulges into the contextualisation of the OAs which more specifically address the issues with students' interest and enjoyment of physics.

Chapter 4

Physics-based Outreach Activities

The physics-based outreach activities that were implemented were developed prior to the onset of this study. This study therefore focuses on their evaluation and on identifying components of an 'effective outreach activity'. The researcher defines an effective OA as one that meets the following criteria:

- Induces positive learning experiences among students
- Successfully applies elements of the compulsory physics curriculum
- Enables students to apply skills of inquiry as well as content knowledge
- Is easily implemented by a teacher and does not conflict with accountability pressures

However the researcher does not rely solely on the instructional design to produce an effective OA. This chapter therefore explains the contextualisation of the OAs and more specifically what was provided to teachers. Extracts of the specific OAs themselves can be found in Appendix 1 and 2.

4.1 Contextualising Physics

The decision to contextualise physics is largely based on the issue that many students find physics boring (see Chapter 2). Lavonen *et al.* (2005) recommend adapting either the content or context of physics in order to make it more interesting and appealing to students. Although they focus on female students, the points they raise are considered to be applicable to all students. Given that the curriculum is fixed and there is no immediate way of changing the specific content that is delivered to students, the aspect that can be changed is the context through which the content is delivered.

The InterAcademies Panel (2010) posits that whether or not students perceive a topic to be worth learning is influenced by the way in which the skills and knowledge they are presented with are applied and put into context. This echoes the findings of Sjøberg (2000) who experimented with delivering the same content to students but through alternative contexts. He found that the context through which scientific ideas are taught were more influential to students' interest than the scientific ideas themselves. This was reaffirmed through the ROSE project findings reported in 2010 by Sjøberg and Schriener.

However, the diverse environments presented by classrooms has been touched upon and therefore finding a context that is well suited to all students and their individual preferences is a trying task. One of the biggest discrepancies among the literature in interest and perceptions is seen between genders. Student accounts often reflect differences between male and females' preferred aspects of science. Typically, girls prefer areas relating to topics such as health, medicine and the human body, and boys prefer topics that are more technical, mechanical and explosive (Sjøberg and Schreiner, 2010; Lavonen *et al.*, 2005; Jones *et al.*, 2000).

However, despite the divergences, we do see some commonalities.

Osborne and Collins highlight one area that presents a valuable "*point of engagement*" (2001, p. 454) and appears to generate "*universal enthusiasm*" (2000, p.5) is topics of astronomy and space. What's more they also emphasise that this topic received interest not only from students who wanted to pursue science but also those who did not. Even students who declared no interest in physics were seen to become animated when entering discussions around space (2001). Similarly, a study conducted by Williams *et al.* (2003) used questionnaires to investigate why year 10 students across state schools (aged 15) in England are not interested in physics. An open question asked students to highlight what they found interesting or boring about physics. Many topics were reported on both sides of the coin, i.e. some students found a topic interesting and others found it boring, however the topic 'Solar System/Universe' only appeared in reports of interesting areas, no students claimed it was boring.

On an international level, data gathered from students across 160 countries and reported by Sjøberg and Schreiner (2010, p. 21) showed clear divergences in preferences across gender, but also found that the most popular topic for both male and female students was '*space, life, wonder and openness*'.

However, astronomy is not a broad area of the compulsory curriculum (Salimpour *et al.* 2018), and therefore it is important to identify how curriculum content can be taught *through* the context of astronomy. Lavonen *et al.* (2005) explain how astronomy can be used to teach a variety of topics in physics, such as kinematics, electromagnetic radiation and nuclear fusion. Trumper (2006b) recognises the cross-curricular links astronomy offers such as mathematics, chemistry and geoscience. This is reinforced by Coward *et al.* (2011) who highlight astronomy's application to both fundamental and applied sciences, and how it clearly translates to professional research, thus providing a relevance and real-life application to students. Astronomy is among the sciences that receive the highest public interest across all countries (Trumper, 2006b). As a subject it spans a broad range of topics and has a wide spectrum of knowledge that provides many applications to the curriculum.

Although complications still remain and topic preferences will understandably vary from studentto-student, the research discussed here indicates the positive potential of astronomy as a context. However, despite the promising potential of astronomy as a context, there are no apparent studies that have specifically investigated its application as a context in the classroom and what influence this has on students' perceptions and learning experiences. This is therefore another gap in the literature that this study addresses.

In response, this study set out to test this potential and evaluate the effectiveness of a series of astronomy-based educational OAs, developed prior to the onset of this study. Astronomy serves purely as a context and the OAs still cover necessary areas of their curriculum, it was therefore anticipated that teachers would not need to find additional teaching time for implementation.

Though, what is evident in the literature is the various outreach initiatives that have provided schools with access to telescopes. Previous and existing educational programmes across the world that have used access to the cosmos as a driving force to promote interest in science. Such endeavours are now discussed before setting out the theoretical framework and methodology of this study.

4.2 Robotic Telescopes

Telescopes come in various types and sizes. They can be sensitive to different wavelengths of light, such as infrared, optical, and x-ray. Arguably the most recognised across the media due to their 'pretty pictures', and those used in this study, are optical telescopes. These observe the Universe in the same way we observe the world around us, through the visible wavelengths of the electromagnetic spectrum.

The key difference between regular telescopes and robotic telescopes (RTs) is the 'hands-on' element. Some students, scout or brownie clubs may have experienced a star-gazing night where they have set up small telescopes one evening and explored the night sky. Users would choose a region of the sky they want to look at and then guide the telescope in that direction. Conversely, RTs do not require this manual control and instead, they are driven, often remotely, by computers.

Remote robotic telescopes, which can be controlled remotely and do not require someone to be on site, are particularly advantageous as they do not demand the observer to be present for their observations. Such telescopes are usually ran by schedulers that programme observation requests submitted by astronomers. This therefore offers greater flexibility in where the telescopes are located. Telescopes are best placed at high altitudes, above much of the atmosphere, weather, and light pollution from surrounding cities. Such locations help to improve the quality of observing and resulting images.

Also benefiters of remote robotic telescopes are schools. No school trips need to be organised anywhere and teachers and students do not need to spend time learning about and setting up a telescope, as they are already set up, aligned and ready to go (Gomez and Fitzgerald, 2017). Students can control a telescope on the other side of the world right from their classroom. This ability in itself could be considered to be an exciting prospect for many students.

We now look to some of the existing educational projects involving robotic telescopes and discuss their levels of success.

4.2.1 Robotic Telescopes in Science Education

There are multiple science education projects around the world based on astronomy and robotic telescopes. A recent paper by Gomez and Fitzgerald (2017) reviews the current educational projects around the world that offer students access to research-grade robotic telescopes. The 24-inch telescope at Mount Wilson is recognised as the first research grade instrument that provided telescope time for educational purposes through the Telescopes in Education project (Gomez and Fitzgerald, 2017; Clark, 1998).

At around the same time in 1993, and more specifically in the U.K., the Bradford Robotic Telescope (BRT) was launched. Astronomers running the telescope encouraged amateur astronomers and members of the public to submit observation requests. Baruch (2015) reported their surprise at the level of interest received from schools and teachers all over the world and soon realised the unique opportunity they were offering to students.

Initially, the BRT targeted primary schools, seeking to work with the youngest students who could use a computer and to provide a window to explore questions that could not be answered in their day-to-day lessons. However, due to such success, the project expanded to secondary schools through to sixth form education. When Baruch explored university applications, a 30% increase was seen in STEM-based subject applications from the 50 schools involved with the BRT compared with 80 schools who were not.

This unexpected success was one of the first pioneers of the exciting and engaging opportunities robotic telescopes have to offer to science education.

Following this, more science education projects incorporating robotic telescopes began to develop in the 1990s. Most involve the use of optical telescopes however projects exist that allow students to observe the Universe through different wavelengths of light, for example Pulses at Parkes uses a radio telescope (Gomez and Fitzgerald, 2017).

The buy-in for using optical telescopes in schools is usually the 'pretty pictures' aspect, where students can choose objects to observe and perform colour imaging processes. However, there is the worry that teachers may consider colour imaging to be the limit to what students can do with robotic telescopes and are unaware of the scientific opportunities they can offer. Access to such instruments provides a unique way for students to practice the 'scientific method', from generating hypotheses, collecting data, performing analysis and drawing conclusions. It is

recognised that the use of remote astronomy in the classroom is slightly more abstract to the typical textbook approach or typical classroom practical experiments. It may also seem very technical and potentially overwhelming to some teachers, consequently deterring them. It is therefore essential to emphasise how these projects tie in to the curriculum content and to provide a high level of support to ensure that teachers feel comfortable and competent in delivering such material.

Where this study implemented and evaluated a series on astronomy-based outreach activities, it was important to first review existing projects in order to identify both their successes and lessons learnt. The researcher purposefully only focuses on research around optical telescopes rather than any other wavelength as it is optical data that is used in one of the OAs of this study.

MicroObservatory is one of the longest standing educational programmes involving online robotic telescopes. The telescopes saw first light in 1995 (Sadler *et al.*, 2001) and with a goal to promote inquiry-driven projects, it encompasses a network of five online telescopes that are accessed by middle and high school students, teachers and the public across the U.S.

In 2007, Gould *et al.* evaluated the educational influence of ten years of the project. With a total of 475 students across the U.S, they assessed students' knowledge gains through pre- and post-knowledge tests and from looking at their project work and also investigated students' interest through an analysis of 2,613 online comments. They were overwhelmed with the levels of enthusiasm and interest from students and report receiving frequent messages from students explaining they had been inspired to pursue astronomy at college (university).

A pilot carried out by Chubb *et al.* (2017) allowed key stage 3 students to take part in the 'Discovery Project' where they were given the chance to discover new variable stars. These researchers gathered feedback from their teachers, who reported that students appeared to enjoy using the activities and software. A particular positive was expressed towards working with real data used in real research. However, they also reported that although initial interest was very high from students, attendance was seen to decline in the middle of the sessions. Overall teachers felt that students enjoyed the variety of software but there were occasions where difficultly levels were perhaps too high.

While it is considered that teacher feedback is important, the researcher believes it cannot supersede the value of feedback directly from students. Concerns in reliability of teachers as sources of evaluation is detailed by Fitzgerald *et al.* (in writing) where it is often seen that teachers generally have a more positive view of the classroom than students or an outside observer, their reports are based on results from multiple studies (e.g. Hook and Rosenshine, 1979; Centra, 1973).

Richwine (2007) ran an astronomy-based science project on secondary school students, aged 16 to 18 that involved studying images of variable stars from an extensive archive. Much of this archived data may never have been analysed by professional astronomers allowing students to conduct never-done-before research.

Interested in the project's impact, Richwine implemented pre- and post- attitude questionnaires and knowledge tests. Although students showed a significant increase in their knowledge test scores, sadly, no significant changes were seen in their attitudes.

However, given the instrument they used, this is perhaps not unsurprising. As most frequently seen in astronomy education research, this study employed Zeilik *et al.*'s 'Survey of Attitudes towards Astronomy' (1997, 1999). This is an adaptation of a previously developed survey by Schau *et al.* (1995) regarding attitudes towards statistics and currently, no study that employs this instrument is known to have found significant changes in a repeated-measures design and is typically seen to be susceptible to ceiling effects (Bartlett *et al.*, 2018).

Although some research has been done to an extent, the majority of telescopes in education projects cease to implement valid and reliable evaluation instruments, therefore suffering from methodological flaws, or, even worse, carry out no evaluation into the effectiveness into their initiatives (Gomez and Fitzgerald, 2017). Although many programmes boast how telescopes *will* change science education, research into the effectiveness of such programmes and how it *actually does* effect student engagement is much more limited. While it is unlikely that such programmes offer nothing to students, in many instances, these do remain as assumptions yet to be properly validated (Fitzgerald *et al.*, 2014; Gomez and Fitzgerald, 2017).

The majority of evaluation is seen to focus on gains in students' content knowledge, and many studies have shown subsequent improvements (e.g. Coward *et al.*, 2011; McKinnon and Danaia, 2008; Gould *et al.*, 2007; Taasoobshirazi *et al.*, 2006). Research into students' attitudes is much more limited and also reports mixed results. Reports are largely anecdotal, involving post hoc data collection, quotes from student and teachers and successful case studies. Although this provides a level of insight into how the programme was received by participants, it provides no indication of the actual impact of the project. This cannot be measured without, at the very least, a baseline understanding of participants' perceptions prior to any initiative or involvement with such a project. Much research shows no significant improvements on students' attitudes towards science at school subsequent to relevant initiatives (Fitzgerald *et al.*, 2015a and Richwine, 2007).

A limitation in evaluation if often seen in accessing participants. Studies are often forced to use astronomy clubs or students that have volunteered to take part in astronomy- and telescope-based activities, therefore already presenting favourable attitudes and are not representative of the secondary school student population (e.g. Coward *et al.*, 2011; Taasoobshirazi *et al.*, 2006).

Fitzgerald *et al.* (2014) point out that the most common reason for carrying out mere surfacelevel evaluation is due to a lack of funding, however they also emphasise that evaluation procedures do not necessarily need to be costly. There are many existing instruments that have been previously validated and can be easily implemented into educational programmes. Just simple pre- and post-questionnaires can reveal a great deal about the experiences of participants and the impact on their attitudes or knowledge.

Some of the major educational projects involving robotic telescopes around the world and the pioneers for such initiatives have been highlighted, and many have been criticised for the evaluation methods they employed. However, there is one project that is of particular interest.

Space to Grow (STG) was an Australian project administered by astronomers from the Physics and Astronomy department at Macquarie University (Danaia *et al.*, 2012). It is of particular interest to this study as not only did it employ both IBL as the pedagogical approach and astronomy as the context, they also provided schools with access to the Faulkes telescopes (see Chapter 4.2.2) via the LCO network in order to obtain real scientific data to use for investigative activities in their classrooms. Though the astronomy topics of their student projects were different to those of this study, there are many comparable elements that remain particularly relevant.

The project was funded for three years (2009-2012) with a target audience of 14-18 year old students in New South Wales, Australia (Danaia *et al.*, 2012). The educational research element of STG was very much a priority in the project. This was designed to investigate the project's impact on students' perceptions of science, their knowledge and their intentions to pursue science post-compulsory education. This was measured with pre- and post- attitude questionnaires, pre- and post-knowledge questionnaires (astronomy-based) and interviews with a sample of participating students (Danaia *et al.*, 2012).

Subsequent to implementation, the researchers saw gains in students' content knowledge, but significant changes were only seen in some aspects of students' experience of science. Students perceived the lessons to be less teacher-led but also felt the project was less relevant to their everyday life than their usual lessons. Aspects regarding their excitement, perceived difficultly level and enjoyment were unchanged. Fitzgerald *et al.* (2015a) suggest that an initiative project with just a few classes a week over a limited time period is not sufficient in overpowering the more common, unexciting day-to-day science lessons students are used to. Such findings could imply that implementation reaped little benefit and is suggestive that the OAs implemented in this study, which are not intended to be extended initiatives, have a poor chance of success. However the researcher emphasises that this study does not seek to induce attitudinal change. Instead it seeks to identify how positive learning experiences are promoted among students and

thus what components should be used in outreach activities and encouraged in physics classrooms.

4.2.2 The Faulkes Telescope Project

The OAs implemented in this study were developed as part of the Faulkes Telescope Project (FTP). FTP was set up in 2004 where concern with students' disengagement with science and maths was even evident then. Dr Martin Faulkes responded to this and sought to make a contribution to science education that would inspire students and improve their attitudes (Beare, 2007; Lewis and Roche, 2009).

With funding from the Dill Faulkes Educational Trust and the UK Government, two 2-metre remote, robotic telescopes were built at observatories in Hawaii (Faulkes Telescope North) and Australia (Faulkes Telescope South). These sites were chosen so that users could observe the night sky in both hemispheres and also because these locations presented time zones that meant they were in the dark whilst UK students were at school, thus allowing for them to carry out real-time observations in their lessons. These were research-grade telescopes that allowed students to explore the wonder and beauty of our Universe, taking, processing and analysing images.

However, FTP set out to provide students with more than just the prospect of taking "pretty pictures" of the cosmos and aim to offer students access to a remote "laboratory", opening the door to the Universe right from their classroom. Teachers can use the telescopes as a tool to carry out real science, whether it relates directly to their school curriculum or form a more extensive research project in collaboration with research members of the FT team (Lewis and Roche, 2009).

In 2005, Las Cumbres Observatory (LCO) bought out a large percentage of the telescopes and the project became an official educational partner of LCO. At this stage, LCO were building a worldwide network comprising of 2-metre (FTs), 1-metre and 0.4-metre telescopes. As of 2018, the network encompasses 21 telescopes at eight sites across the globe, Australia, China, Israel, South Africa, Tenerife, Chile, Texas and Hawaii. These locations create a unique network where at least one site is in darkness throughout a 24-hour period, allowing for continuous observing. Any school that signs up with the FTP will have access to each of these telescopes.

Schools throughout the UK and Europe have embarked on various scientific adventures through this opportunity. Projects have included collaborations of schools in different countries where they have shared and combined data, thus presenting the additional opportunity of working with students overseas and developing communication and teamwork skills.

An example of this is shown in Figure 4. This displays a light curve of asteroid (1676) Kariba produced as a collaboration from schools from five different countries, Poland, UK, France,

Portugal and Ireland. Each school observed the asteroid for a given time slot and then they combined their observations and analysis to produce this extended light curve.





There have been several occasions where schools have been acknowledged in published journal articles for their contribution of observations and analysis (Lewis and Roche, 2009), these are clearly big success stories, however again, other than anecdotal reports, no real investigation has been carried out into the impact such involvement had on students' perceptions of and engagement with science. What's more, despite the potential links to the curriculum, the majority of FT work in schools has been as enrichment activities rather than curriculum specific.

A degree of evaluation of the impact of the FTP was carried out by Beare (2007) when the project was still starting up. Unable to access student participants, the research was carried out with 15 trainee primary school teachers. All participants had studied biology up to A-level but only one had studied physical science; Beare therefore claimed them to be comparable to key stage 4 students.

This research focused on perceptions towards directly controlling the telescopes, performing observations and carrying out various astronomy projects relating to an astrophysics option in an A-level physics course. Responses to Likert-scale questionnaires, although not validated, revealed positive attitudes towards the telescopes, encouraging excitement and inspiration.

The results spurred confidence in the success of the project, suggesting it could be very valuable in promoting students' enthusiasm in physics. However, since this small-scale research project, no further investigation has been carried out on the Faulkes Telescope Project until now.

This study therefore has not only carried out thorough evaluation of outreach activities from the Faulkes Telescope Project, but also how FTP is not only useful for enrichment but can also be applied to the compulsory curriculum content, and potentially feed into teachers' schemes of work.

4.3 Down2Earth

Down2Earth is a separate educational project ran by FTP in partnership with the National Museum of Wales. With supporting funding from the Science and Technology Facilities Council (STFC), the project offers a series of teacher and student resources based around asteroid, comets, meteorites, impacts and mass extinction.

One resource is the Impact Calculator, developed by E. Gomez and J. Yardley and based on the original online tool, Earth Impact Effects Program (Collins *et al.*, 2005). The Impact Calculator is an online simulator that allows students to design a virtual asteroid or comet and simulate its collision with Earth. Using a Google Maps interface, students can choose a region on the globe that they want to strike and a scientifically-sound impact crater is superimposed onto the Earth's surface. This resource was developed to appeal to both students and larger audiences (Astronomy and Geophysics, 2008) and can be used as an inquiry-based tool that can apply various aspects of the science curriculum.

To date, no formal evaluation of this online simulator as an educational tool has been carried out. This is therefore another first for this study as some of the OAs that are evaluated are lessons built around the Impact Calculator. Although this study presents the first formal evaluation of this particular simulation, research into other computer simulations has been carried out before.

4.3.1 Computer Simulations in Science Education

Investigations into the influence of computer simulations as an educational tool has been explored quite heavily in the past few decades as technology has developed and schools' access to technology has increased dramatically.

Many schools are faced with a lack of funding to support practical tasks in science classrooms but this can be overcome through e-learning and free online resources that simulate laboratory practices (Agrusti, 2013). Computer simulations are used to model different processes within a real system. They are useful in education as they overcome the expense of equipment for experiments and do not require time to set up. The National Research Council (1996) also highlight their ability to match the varying pace across different students and can therefore account for differentiation in the classroom, an issue highlighted by Ofsted (2013) that many schools are not accounting for.

The dramatic advancement in technology over the past few decades, has not only led to significant developments across industries and changes to lifestyles, but it has also fed into educational reforms (Bernard, 2017). As of 2014, 70% of primary and secondary schools in the UK had access to tablets (Coughlan, 2014).

However, such advances in technology and the access for schools are seemingly not being utilised to full potential. Webb (2010) argues that even where computers have been incorporated into science lessons, they are often merely used as a supportive tool for traditional, teacher-led pedagogies such as animations or demonstrations. Nonetheless, there are circumstances where computer simulations have been implemented more constructively and have also been evaluated.

Such research typically focuses on the effectiveness of computer simulations (CS) compared to hands-on laboratories (HOL), remote laboratories, traditional teaching methods, or a combination of the three. There are also some studies that focus on CS as entire replacements of traditional instruction and some as enhancers of it (Rutten *et al.*, 2012). Despite the abundance of research, the debate remains long-standing and there does not appear to be a clear-cut winner (Ma and Nickerson, 2006; Smetana and Bell, 2012). The researcher suspects that this is largely due to the variety of such simulations and research rapidly becoming outdated as technology accelerates its advantages. Corter *et al.* (2004) argue that the majority of comparative studies demonstrate simulations to be a worthy substitute for hands-on labs. However as with the issue of IBL, review of the literature demonstrated that many studies have focused on undergraduate courses rather than secondary schools.

Much evaluation and educational research has focused on the benefit of technology on learning and engagement. In light of the broader research discussed in Chapter 2 that reveals the importance of attitudes in education, it is necessary to consider the opinions among students towards the use of computer-based learning and whether they reap the benefits. Though this study focuses on learning experiences, research into knowledge gains is also discussed here as this is ultimately what the teacher will be interested in when considering implementation. The researcher considers both students and teachers views to be influential on *if* and *how* the computer-based OAs are adopted in the classroom (Li, 2007).

Sarabando *et al.* (2014) investigated 12-13 year old students across three different scenarios: just hands-on activities, just a computer simulation and both methods combined. Participants involved three teachers and six classes (each teacher had two classes). Students' knowledge gains were measured through pre- and post-tests, the results of which are summarised in Table 4.

Table 4 – Results from Sarabando <i>et al.</i> (2014, p. 119)			
Teacher	Lesson Implementation	Total Gains	
Teacher A	Class X (HOL)	20.0%	
	Class Y (HOL + CS)	40.7%	
Teacher B	Class X (HOL)	29.6%	
	Class Y (CS)	58.4%	
Teacher C	Class X (HOL)	36.9%	
	Class Y (CS)	45.0%	

It was evident from the results that in all scenarios involving CS, students had higher knowledge gains. Interestingly the one class that used both HOL and CS (teacher A, class Y), achieved the smallest gains out of the other class Ys which used just CS. Though, teacher A also had the smallest gains for just the HOL of the three teachers so it is likely that an element of teaching style is at play here. Each teacher was also from a different school and therefore introduces additional variables that cannot be accounted for.

One divergence among the research is seen where some studies focus on comparing CS to traditional instruction, others evaluate its ability to enhance traditional instruction. It should be emphasised that here, the researcher did not set out with an aim to produce evidence in favour of *replacing* hands-on experiments with CS. As with other researchers, it is perceived that real, hands-on, laboratory experiences are an essential part of education (e.g. Ma and Nickerson, 2006). Instead, in instances where hands-on labs are not feasible, an OA involving a computer simulation could offer an alternative way of enabling students to engage in experiences of investigation and inquiry. The literature reviewed here is intended to demonstrate that such feasibility is not at the expense of students' learning experiences and outcomes.

For example, de Jong *et al.* (2013) highlight how real hands-on labs can be limited by practical and financial constraints. Not only are HOL costly, they are also space and time consuming (Corter *et al.*, 2004; 2007; Sauter *et al.*, 2013), which have been highlighted by teachers as barriers to their teaching. CS therefore enable students to gain experience in scientific experimentation that may otherwise be unattainable to schools through more traditional means (Sauter *et al.*, 2013).

Less specific to simulations but still considered to be relevant are the results from Li (2007). Questionnaires were implemented to 575 students ages 12 to 18 about their perceptions regarding technology integration in school. They identified areas relating to confidence to be most frequently discussed, relating to 74% of students comments. Typical comments reflected on the increased access to information and research that technology provides and how it makes learning easier. The second most cited topic related to pedagogy, students were particularly fond of the visual and animation elements that technology provides and how it can bring the real world into the classroom.

Webb (2010) emphasise a unique selling point of CS in that they can be used to explore scientific phenomena that are too small, too large, too slow or too fast to observe in real life. Examples of this could include the evolution of a star, nuclear decay or global warming. In the case of this study, it is the process of an extra-terrestrial object colliding with Earth's surface (see Chapter 4.4.1). This is not an event that occurs frequently and is not a phenomenon you can get up close to when it does. Being able to simulate such events therefore provides unique opportunities for scientific learning. Huppert et al. (2002) are also of the opinion that CS enables students to carry out investigations quicker, therefore providing them with more time to generate hypotheses and analyse results. In specific relation to IBL, Mäeots et al., (2008) carried out a study on 302, 10 to 12 year old students in Estonia that looked specifically into the inquiry element of simulations. Students were given pre- and post-tests that assessed their level of inquiry skills. Generally, students' initial inquiry skills (pre-test) were relatively low but results showed significant developments, subsequent to a web-based learning environment initiative. The largest improvements were seen in problem identification, formulating research questions and hypotheses, and applying skills into data analysis. Similar findings were reported by Nakhleh et al. (2000) in open questions asking students to report the advantages and disadvantages of using computers. These revealed several advantages of computer labs, this included the ability to cover more content over a shorter period (thus reflective of Huppert et al.'s argument) and being able to work at their own pace. This final point is reflective of that made by the National Research Council (1996).

The effectiveness of CS however is argued by Grant (1995) who claims that simulated labs do not allow students to learn through trial and error, however here the researcher considers this to be largely untrue. Although this argument may have been true twenty years ago, is likely to have become obsolete given the advancement in technology over this time. For example, the 'Recreate a Crater' OA implemented in this study (see Appendix 1) that uses a relatively basic simulator, involves students designing an impactor that will recreate one of the existing craters on Earth. Students have to experiment with changing various parameters that have different effects on the object's kinetic energy and damage to Earth, until they succeed. Such a process is considered by the researcher to be a classic example of trial and error.

Corter *et al.* (2007), also found CS to be at least equally effective (if not more) as HOL in developing learning and understanding, they also suggested this was due to the additional time provided in CS to develop their conceptual understanding that is lost in HOL through assembling and taking down equipment.

Evidence is therefore present within the literature for computer simulations to be effective in promoting student learning and positive experiences. Some research is even indicative that it offers additional opportunities and overcomes some of the boundaries of HOL. The researcher therefore considers the OAs based around the Impact Calculator simulation not only to have potential to promote positive learning experiences among students but also appeal to teachers due to the learning opportunities they offer.

4.4 Educational Outreach Activities

Chapter 3.3 highlighted some of the barriers that teachers face and can often lead them to be hesitant towards implementing inquiry-based approaches. Two particular barriers were time and knowledge limitations. The OAs produced in this study therefore aimed to eradicate such concerns.

Fitzgerald *et al.* (2015b) highlight two simplistic techniques of gaining the approval of teachers and counteracting their resistance. One is to provide as much of the required material as possible, therefore each activity was provided as a "Teacher's Package" that included the following:

- A lesson plan with suggestions and guidance on implementation for the teacher
- A series of *background information* relevant to the subject matter, this is so that teachers will not need to do any background research into the context in order to deliver the activity
- *Student worksheets*, these are for students to record their workings in throughout the investigation activity
- *A teacher guide*, for guidelines of answers that students should be getting and suggestions for discussion points
- *Introductory PowerPoint presentation*, for if the teacher wants to deliver a short presentation at the beginning of the lesson to introduce students to the subject they will be investigating

The only equipment that will need to be provided by the schools themselves is computers, laptops or tablets, which as identified by Coughlan (2014) are available in the majority of schools.

The second technique to appeal to teachers is to make the OAs editable so that teachers are able to customise content in order to make it most suitable for them and their students. It has been touched upon and will be discussed in detail in Chapter 6, that no two classrooms are identical, therefore the researcher did not assume that one approach or method would be optimal for all classrooms. Although it has been argued that the instructional design of the OAs will be through structured inquiry, how exactly the activities are implemented is largely at the discretion of the teachers. Teachers were provided with multiple suggestions and options of how the activity can

be carried out, enabling the teacher to take a direction in which they are most comfortable with. Editable formats are also useful in differentiating lessons as it means the teacher can adjust activities to best suit the varying needs of their individual students. Techniques for differentiation were provided in the lesson plan and the researcher also explained possible variations to the teacher ahead of implementation. Having said this, it was emphasised to teachers that the OAs are designed so that learning is student-centred, and they were encouraged not to alter the activities too much so that the focus of the lesson shifted back to the teacher.

The activities were also developed in a way that ensures students are not all doing the same thing, work can be divided among them, and students can collaborate, share their findings and discuss their work with peers. This decision was influenced by the findings by Treisman (1992) who compared the performance levels of students who studied individually and students who spent at least a third of their time in groups, discussing their work. Not only was it found that those who studied in groups performed much better, but after encouraging the students who worked individually to then work together, their performance levels were also seen to increase.

However, it was not only important to consider the needs of students and teachers but was essential that the OAs were accessible to schools more generally (Fitzgerald *et al.*, 2015b). The basic versions of the activities use Microsoft Office (e.g. Word, Excel, PowerPoint) so were not anticipated to be a problem for schools as they do not need to download any specialist software. Students though needed to access particular websites, however these were not expected to be blocked by any firewalls put in place by school internet servers and had not presented any problems in schools previously and external to this study.

The next two subchapters will now describe the specific OAs that were implemented in secondary schools in this study. Activities were designed for key stage 3 and 5 and all apply a context of astronomy and space science while focusing on the physics content within the national curriculum for England and Wales. OAs for key stage 4 students were also available but implementation of these did not take place throughout the duration of the study and therefore are not reported.

4.4.1 KS3: The Impact Calculator

Activities developed for key stage 3 students (years 8 and 9) were based on the Down2Earth Impact Calculator. Students investigate various comet and asteroid impact events and their effects on Earth. Through inquiry-based processes, students explore the outcomes of changing the properties of an impactor. The activities cover aspects within the physics curriculum such as speed, distance and time; kinematics and energy.

Most students at this stage of education will not have come across the equation for kinetic energy, this therefore gives students the opportunity to first explore the different parameters and relationships rather than just being presented with the facts and underlying equations. Not only do the activities cover compulsory curriculum content and therefore do not defy accountability measures, but they are also intended to drive students to generate hypotheses, identify their variables, record their results and interpret their findings. This is therefore considered to be reflective of constructivism where students learn through their own investigative experiences.

Several activities had been developed around the Impact Calculator which teachers could choose from, extracts from one example, 'Recreate a Crater' is provided in Appendix 1.

4.4.2 KS5: Hubble's Expansion of the Universe

This teacher package is designed for sixth form or college students. Although it is recognised that these students have already chosen to pursue physics in post-compulsory education, the researcher considers the importance of reinforcing this interest and continuing encouragement through to university.

The activity involves students working with real photometric data of type Ia supernovae which are used as standard candles in distance measurements in cosmology. By measuring how bright they appear from Earth (their apparent magnitude), we can compare this with their known absolute magnitude and assume that the difference between these two values it due to the supernova's distance from Earth.

Once students obtain their value for distance, they access an online astronomical database to search for the radial velocity of the host galaxies of the supernovae. They can then produce a Hubble plot of distance against radial velocity for 15 supernova datasets. Using the equation y = mx + c (that students should be familiar with at this stage of their education), students calculate the gradient of their graph to obtain a value for the Hubble constant (H₀). They can then use this to derive an approximate age of the Universe.

This activity allows students to work from the beginning to see how Hubble originally created his first Hubble plot. Using real data they have analysed themselves they are able to calculate the Universe expansion rate and its approximate age using real data directly from the classroom.

This activity can be differentiated depending on the ability of students and also how much time the teacher wants to devote to this area or how much the students want to take on in their own time, though can be done over a single lesson. An extract of the activity is provided in Appendix 2.

Again, this activity covers additional skills beyond the curriculum content. Students gain experience with accessing research databases to extract data, plotting graphs, extrapolating data, unit conversions and advanced calculations, thus providing skills not only useful for their exams but also useful for pursuing physics, astronomy or another science at university level.

4.5 Summary

This chapter has identified a key gap within the literature whereby although there is much evidence in support for astronomy as a point of engagement, there is a gap in the research in terms of how this is applied to practice and can influence students' learning experience in compulsory physics.

This chapter has explained the context through which the OAs will deliver physics curriculum content to students. It has also set out the format of the OAs for implementation across secondary schools in England and Wales. The OAs were designed so that learning is student-centred and students are able to take control of their learning and carry out investigations into themes of astronomy that have been seen to trigger interest amongst various groups of students. The OAs apply compulsory curriculum content that will have allocated lesson time and therefore should not be considered as enrichment activities that demand additional teaching time.

Attempts to overcome barriers to implementation include providing OAs that can be implemented over a single lesson, requiring minimal equipment and providing teachers with all the material necessary so that the only work demanded of them is to read the activities. All material was also provided in editable formats so teachers could make changes and adaptations to suit their students if they so desire.

Specific extracts of the OAs are provided in Appendix 1 and 2.

Chapter 5

Evaluation Framework

Though the evaluation framework was not fully conceptualised until much of the data had already been collected, it is described here in order to illustrate the underlying structure and theories that set out the foundations for the specific research design and methodology. The conception of the evaluation framework emerged largely as a result of the researcher's reflexivity, as her involvements in and understanding of the classroom environment and students' learning experiences developed and thus influenced her interpretations of the research environment and evidence that came to light.

Stake (1995) explains how at the beginning of his inquiries, he sets himself a page limit for his final report. He justifies this technique in that it prevents him from continuing to gather more and more data on the possibility it may be of interest. Application of such restrictions forces him to prioritise data of value. A similar intention is a primary purpose of the evaluation framework of this study. The constructivist, inductive approach of this study that applied Scriven's (1991) goal-free evaluation, yielded extensive amounts of data that needed to be refined and prioritised. As Stake warns, much of the data was considered to be of potential interest but would deviate from the focus of this study.

The framework therefore sets out the structure for what is reported for each case in this study. It prescribes a set number of areas of discussion and anything outside such dimensions is deemed beyond the scope of this particular study.

Figure 5 illustrates the mind-map that evolved as qualitative data was thematically analysed and patterns were identified. This image is a reconstructed version of a hand-written mind-map. It has been transferred into a digital format in order to make the text clearer and to ease reading. Table 5 details a key of the components of Figure 5.

This chapter therefore explores the various theories that underpin the themes identified among the data collected and that are considered to be key in promoting positive learning experiences for students. Such processes were critical in the researcher's reflexivity processes and thus the development of the evaluation framework.

Learning experiences, engagement, motivation, interests and confidence among students are not straight forward, one dimensional concepts. They are largely multifaceted, latent constructs that have welcomed a wide array of theories. This chapter consequently reviews various theories surrounding such relevant concepts in order to construct a framework that will form the basis of the methodology and data analysis of this study.

Table 5 – Key of Evaluation Framework Mind Map			
Item	Explanation		
Large blue squares	These represent the research questions of the study and were used to identify what data evidence they related to.		
Grey squares	These represent some of the key messages that were identified among the data relating to students' learning experiences. More specifically, they summarise themes that were commonly identified across both statistical and thematic analysis (see Chapter 9) and were therefore considered to be important factors.		
Small rectangles	These represent overarching dimensions that were considered to tie the themes among the data to the research questions. Such dimensions were also situated within the existing literature. They were therefore recorded as potential dimensions of the framework.		
Arrows and text	Arrows indicate that relationships were identified between two or more aspects. Free-written text provides a brief explanation for such relationships and what this may imply for the evaluation framework and how it feeds into the grey squares (key messages).		


Figure 5 – Mind-Map Used to Conceptualise the Evaluation Framework

5.1 The Concept of Interest

Chapter 2 revealed not only the negative impact resulting from students' lack of interest in physics but also the power interest can have on students' persistence in physics. Though there are multiple problems within school physics yet to be successfully addressed and thus form some of the targets of the OAs in this study, the researcher considers that students' interest should be triggered from the introduction of the OA in order for them to consider it worthy of their effort. We therefore begin by addressing theories of interest in order to define the term and identify methods of enhancing and fostering interest.

In terms of definition, the literature largely focuses on two categories, individual interest and situational interest (Schiefele, 1991; Krapp *et al.*, 1992). Individual interest is specific to an individual and is generally considered to be a temperament that is acquired with time. This form of interest is stable and is typically associated with positive emotions and raised knowledge. Situational interest is typically more short-term and is triggered by a particular experience or environment; it is therefore generally not unique to but common across individuals.

Though only short-term, Hidi and Renninger (2006) emphasise that situational interest can stand as a foundation from which a more stable, individual interest can develop. They propose a fourphase model of development. Phase one involves triggering students' situational interest, this is then maintained in phase two, an individual interest begins to emerge in phase three and a welldeveloped individual interest is present in phase four. Phases one and two are largely governed by external factors, whereas in phases three and four, interest is more self-generated.

Despite the different definitions, both individual and situational interest involves experiences and psychological states in an individual that are largely referred to as interest. These include increased attention, greater concentration, enjoyable feelings, applied effort and an increased willingness to learn (Krapp *et al.*, 1992). However, given the nature of the two definitions and the nature of this study, the focus here is appropriately situational interest. The embedded, longstanding nature of individual interest makes it an unrealistic target for this study. Instead, the researcher seeks to provide OAs that can trigger students' situational interest, thus addressing phase one of Hidi and Renninger's model.

Ainley *et al.* (2002) acknowledge the abundance of researchers who maintain that bases for situational interest are especially valuable among students who hold no pre-existing individual interest in a subject (e.g. Hidi, 1990; Hidi and Harackiewicz, 2000), therefore placing situational interest as a suitable target for those students who do not engage with physics lessons. Given that Chapter 2.5 highlighted how many students enter secondary education with favourable attitudes towards science that go into decline with progression through education, the researcher considers that whereby an interest once existed, it should be possible to trigger this interest again under suitable situational contexts.

Various researchers emphasise a number of different methods to trigger situational interest, though here the research pays particular attention to the examples provided by Hidi (1990) that include novelty, surprise and unexpected events or ideas. Given that many students perceive physics to be uninteresting or boring (Ametller and Ryder, 2014; Williams *et al.* 2003), which is considered to be based on their relatively prolonged experience of the subject in secondary education, the researcher hypothesises that if students are presented with a novel, unexpected context in their physics lesson and an activity that they would not expect to do in their day-to-day classroom, this deviation from the norm in conjunction with a promising context (as highlighted in Chapter 4), could offer an effective trigger.

Context is also a key element in constructivism in that it assumes that knowledge transfer is facilitated when students are involved in authentic tasks that are anchored into meaningful contexts and the experiences of the learner within such contexts (Ertmer and Newby, 2013).

5.2 Self-Determination Theory

Self-determination theory (SDT) is proposed by Gagne and Deci (2005) and posits that individuals have three basic needs that must be fulfilled in order for one to feel intrinsically motivated. These are competence, relatedness and autonomy.

Competence is experienced when one feels capable and effective within an environment and is therefore considered to tie in with students' sense of confidence. SDT posits that motivation is encouraged when one feels competent and receives positive feedback from a significant individual that indicates such competence. A classroom environment is optimal under circumstances whereby the teacher outwardly conveys their confidence in their students' abilities and also guides them towards a higher level of achievement (Stipek, 2002). This latter point in particular places further emphasis on the importance of successful differentiation among students.

Relatedness ties in with concepts of community and belongingness discussed by Dewey (1958) and Vygotsky (1978). Dewey in particular poses that students should function as a social group as it is through collaboration between peers that learning takes place. SDT poses that students experience relatedness when they feel connected with other individuals in their environment, in the classroom this would be their peers and the teacher. When this personal relatedness is felt by all members of a class, a community is considered to exist (McMillan and Chavis, 1986). Osterman (2000) also emphasises a link between these basic needs and explains that when students perceive a sense of belonging within a group they also begin to feel more competent, both of which enhance intrinsic motivation. Johnson and Johnson (1999) emphasise how cooperative work can not only build and maintain relationships among peers but also establish them. As these relationships grow, improvements are also seen in student moral, commitment and willingness to persist with set tasks.

Autonomy is encountered when students experience an element of control over their learning (Gagne and Deci, 2005). There is also evidence among the research that suggests this is something students would like to see more of in their physics lessons (Angell *et al.*, 2004). Students are seen to be more committed to learning and applying themselves when teachers show an interest in their opinions (Furrer *et al.*, 2014). Angell *et al.* (2004) however report that both teachers and students indicate that suggestions from students are rarely used to guide instruction.

Given that one of the key elements of constructivism is a shift in focus from teaching to learning (Ertmer and Newby, 2013) and thus the classroom becomes student-centred rather than teachercentred, autonomous learning should therefore be clearly evident among the OAs both to address students' basic needs and to also coincide with the emphasis on the control of the learner as outlined by their instructional design.

5.3 Individuals' Internal Goals

The researcher considers that the competence element of SDT is also influenced by students' goal orientations in the classroom. Dweck (1986) highlights how students will generally display one of two goal orientations, performance goals or learning goals. Those who identify with the former are of the belief that the pursuit of tasks is governed by their ability level. Those who identify with the latter however believe that task pursuit is governed by the effort they put in. There is a large limitation of those with performance goals in that they generally stay within their comfort zone and do not seek challenging tasks as they believe their ability is fixed. They must be confident when pursuing a task otherwise they will avoid it. As a result, their learning is limited and these students struggle to expand their ability.

Students with learning goals however, will often seek a challenging task as they believe that if they apply effort, they can succeed. This is even evident among those of lower ability as they believe that choosing a challenging task will foster their learning (Dweck, 1986). These types of students will also respond to failure in alternative ways. Where performance students will attribute failure to a lack of ability, learner students will alternatively respond by increasing their effort or adjusting their approach to a task. Students with learning goals therefore have a less restricted view of their abilities. More recently Dweck (2016) has coined this theory 'growth mindset'.

It is therefore considered that lessons should focus on qualitative goals that are descriptive rather than quantitative. Quantitative goals are more typical of educational grades that students are familiar with. Qualitative goals however offer a less competitive target for students and are less comparable between peers. Such a nature can potentially drive students to aim to learn something rather than obtain a particular grade.

5.4 The Teacher's Role in the Classroom

Despite the instructional design of the OAs that places students and the centre of learning, the teacher still has a critical role (Reigeluth, 1989). Various terms are used in the literature to describe the teacher's role, such as a guider or facilitator (Crawford, 2007). Regardless of the term, the level of support or involvement provided by the teacher is primarily what distinguishes the different levels of inquiry discussed in Chapter 3, and therefore too little or too much input from the teacher can have a negative impact on students' learning experiences.

The teacher is also the medium of instruction in the classroom, governing lesson tasks and activities. It was therefore considered that the teacher will ultimately have a significant influence on students' learning experiences, regardless of the OA. That is, even if two classes completed an identical activity, they could have different experiences based on how their teacher interprets and adapts the activity for their delivery.

Teaching style and approach to instruction is also not the only influence the teacher has on students' experiences. Teacher-student relationships are also seen to be a significant factor associated with student engagement (Roorda *et al.*, 2011; Pianta *et al.* 2012).

Pianta *et al.* (2012) emphasise that teachers should be sensitive to their students, whereby they are aware of and responsive to students' individual needs, and also show regard to students' perspectives. This former aspect is again reflective of a constructivist teacher, where students' perspectives are acknowledged the teacher provides opportunities for students to take the lead and have a formative role in the lesson content (Pianta *et al.*, 2012).

The teacher is also highly involved in self-determination theory. Teachers can promote students' feelings of competence where they express confidence in their abilities and guide them towards higher levels of achievement (Stipek, 2002). Autonomy is enhanced when teachers attribute a level of control to students, and relatedness not only includes a sense of belongingness among peers but is also improved when teachers show respect and warmth towards their students (Furrer *et al.*, 2014).

5.5 Constructing the Framework

All of the theories discussed thus far provide a foundation for understanding influential factors to students' learning experiences in the classroom. Many theories consist of complementary factors and are also seen to be correlational. For example, confidence and interest both influence students' engagement and learning experiences individually, but students with high confidence in a subject are also generally seen to have a higher interest in that subject (Barmby and Defty, 2006; Williams *et al.* 2003).

Having identified the key theories underpinning students' learning experiences, the final process before constructing a refined framework involved identifying existing frameworks within the literature that have been used to evaluate educational OAs from a similar perspective.

What was immediately evident from this search was the wide inconsistency among the terminology for what is described here as an outreach activity. It was apparent that not only were equivalent elements termed differently, but there were also instances where the same terms were used for different things. This not only created difficulties in performing the literature review but was also indicative of the disparate research into this domain. The researcher continues to attribute the term 'outreach activity' to an activity (paper or electronically based) that can be used within a lesson to teach or learn aspects of the necessary curriculum.

It was also evident from the search that much of the literature focused on the evaluation of webor digital-based OAs. Although the OAs developed in this study will incorporate digital elements, this study is concerned with the pedagogical and instructional components of the OAs rather than technical and functionality components. However, many of the dimensions set out by such research were considered to be applicable to both paper- and digital-based OAs as they outline the technical as well as the pedagogical usability, therefore the researcher focused on the latter dimensions. Six studies were identified to possess relevant dimensions in their frameworks; the authors and dimensions are summarised in Table 6.

It is evident from Table 6 that there are several commonalities across the existing frameworks which have been highlighted accordingly (green). Such commonalities are suggestive of the importance of these dimensions as essential characteristics of an educational OA. Hadjerrouit's (2010) framework is considered to be the most applicable to this study as the majority of his dimensions focus on pedagogical aspects rather than technical, and many of such dimensions are relevant to the theories that have been discussed in this chapter to underpin students' learning experiences.

The researcher therefore adopts those dimensions that harmonise the various theories explaining students' learning experiences most proficiently and is also essential to an inquiry-based, constructivist approach to learning. The following subsections therefore discuss each dimension within the framework of this study.

Reeves (1994)	Quinn (1996) adapted by Albion (1999)	Squires and Preece (1999)	Hadjerrouit (20
Epistemology	Clear Goals and Objectives	Match between Designer and Learner Models	Understandability
Pedagogical Philosophy	Context Meaningful to Domain and Learner	Appropriate Levels of Learner Control	Added Value
Underlying Psychology	Content Clearly and Multiply Represented and Multiply Navigable	Prevention of Peripheral Cognitive Errors	Goal-Orientation
Goal Orientation	Activities Scaffolded	Understandable and Meaningful Symbolic Representation	Time
Experiential Value	Elicit Learner Understandings	Support Personally Significant approaches to Learning	Interactivity
Teacher Role	Formative Evaluation	Cognitive Error Recognition	Multimedia
Program Flexibility	Performance should be 'Criteria- Referenced'	Match with the Curriculum	Motivation
Value of Errors	Support for Transference and Acquiring 'Self-Learning' Skills		Differentiation
Motivation	Support for Collaborative Learning*		Flexibility
Accommodation of Individual Differences	, C		Autonomy
Learner Control			Collaboration
User Activity	-		Variation
Cooperative Learning			
Cultural Sensitivity	*Albion only		

Table 6 – Existing Evaluation Frameworks within the Literature

5.5.1 Goal Orientation

The Goal Orientation dimension was adopted from Reeves (1994), Nokelainen (2006) and Hadjerrouit (2010), though is also similar to the Objectives dimension set out by Quinn (1996) and Albion (1999). These authors consider goals and objectives to be vital elements of a lesson in order for students to understand what they are aiming for and what they should set out to accomplish in order to focus their attention (Quinn, 1996).

Goal Orientation is also relevant to the theory that describes the type of goals students approach a lesson with relating to Dweck's growth mindset theory. This dimension is therefore used in order to assess whether the OAs can successfully direct students towards a more learning focused orientation which is advantageous over a performance orientation.

Reeves (1994) and Nokelainen (2006) however focus more on the clarity of lesson goals and both propose a continuum of its specificity. Reeves proposes a spectrum spanning from sharply focused to unfocused, whereas Nokelainen suggests a range from concrete to abstract. Where a specific goal is positioned on either of these continuums is, to an extent, dependent on the subject. For example, where mathematics is a highly prescriptive subject a lesson goal may be to understand and apply simple quadratic equations. However, in an art lesson, the goal may be less prescriptive, for example to gain an appreciation for the work produced in the renaissance era. Reeves' spectrum however is considered to be most applicable to science education.

Here the researcher employs two sub-dimensions of Goal Orientation, one ranging from the type of objective employed by the student, performance to learner and also Reeves' continuum of the clarity of the objectives presented in the OA.

5.5.2 Differentiation

Differentiation is a vital feature of lesson plans and classroom conduct (Ofsted, 2013). It involves constructing the lesson content and information so that it is accessible to all students within a class and can account for individual characteristics, such as ability, gender, language and prior knowledge and experience (Hadjerrout, 2010). Nokelainen (2006) use the term flexibility which is considered to be parallel to differentiation.

The differentiation dimension is considered to hold relevance to a number of theories underpinning students' learning experiences. For one, effective differentiation is deemed important in enhancing students' sense of competence set out in self-determination theory and confidence described in Chapter 2. By tailoring a lesson or activity to the individual needs of students, it is considered that they are less likely to be faced with struggle or failure and are subsequently able to thrive in the lesson. Differentiation also ties in with the Goal Orientation dimension in that all students should be presented with a goal that is suitable for them, that is realistic but challenges them.

Ofsted (2013) found that in key stage 4, the majority of teachers were relying on student sets to differentiate by ability and were not taking suitable measures to accommodate for within class differences. However, in instances where Ofsted did see teachers accounting for within class differences, effective differentiation was seen to be the most effective method of raising achievement in science.

This dimension therefore assessed how effectively the OAs were able to accommodate for differentiation both between classes and between individuals. Though Reeves (1994) proposes continuums within his framework, Differentiation is not a dimension he included. The researcher therefore proposes a continuum that spans from unaccommodating, denoting that needs for the majority of students within a class are not met, to accommodating, where all students' needs are sufficiently met.

5.5.3 Student Autonomy

Student Autonomy is a dimension coined by Hadjerrouit (2010) though is also used in the frameworks of Reeves (1994) Squires and Preece (1999), and Nokelainen's (2006) under the term Learner Control. Student autonomy is an important aspect of inquiry-based learning within the constructivist domain and also one of the three basic needs according to SDT (Gagne and Deci, 2005). Ertmer and Newby (2013) highlight the shift in focus from teaching to learning and that there is an emphasis on control of the learner with this approach.

This dimension therefore assesses how much control is in the hands of the students. Hadjerrouit (2010) describes autonomy as the ability for students to acquire relevant knowledge, independently and without sole reliance on their teacher. The continuum set out by Reeves (1994) ranges from non-existent to unrestricted however given that all students will experience the same OA, variations in autonomy will depend largely on implementation style by their teacher. The researcher therefore wanted the teacher's role to be more evident in the description, and so the continuum instead was chosen to range from dependent to autonomous. Where dependent signifies that students are entirely reliant on the teacher for their learning and autonomous indicates that learning is entirely student-centred.

5.5.4 Collaborative/Cooperative Learning

Classrooms, by design, are social environments. In every classroom that a student experiences, they will have social interactions with the teacher, their friends and all other students whom they may not have friendships with (Urdan and Schoenfelder, 2006). School is not merely an environment whereby students seek to meet performance and achievement needs, but also seek needs of membership and involvement (Gagne and Deci, 2005; McMillan and Chavis, 1986). The value of Collaborative/Cooperative (C/C) Learning was first emphasised by the highly influential educationalists in the early 1900s, John Dewey and Lev Vygotsky. Although they had different views

of education and learning, both conceptualised learning as a social process. Dewey (1958) in particular emphasised that it is through collaboration that learning occurs (Osterman, 2000).

This is another dimension adopted from multiple authors, Reeves (1994), Quinn (1996), Albion (1999), Hadjerrouit (2010) and Nokelainen (2006), and is described by the latter two authors as a process whereby students work together in order to reach a common goal. The researcher also considers that students who have highly performance oriented goals are likely to be more competitive than those with learning oriented goals, and will also more likely compare themselves to their peers. However, students working either cooperatively or collaboratively is deemed to be a more supportive environment where students are focused on working together towards a common goal. Johnson and Johnson (1999) also highlight that cooperative group work can help students to establish and maintain relationships, which brings us back to self-determination theory and the need for relatedness.

The researcher acknowledges an importance in clarifying the difference between cooperative learning and collaborative learning. Here, the researcher regards the former definition as an environment where students help one another and form a supportive environment. The latter definition is denoted by students specifically working together to reach a common goal.

It is considered that where a classroom activity encourages or even requires students to work collaboratively or cooperatively, students' sense of belonging is fostered or enhanced and the classroom holds a greater sense of community (McMillan and Chavis, 1986; Osterman, 2000). Juvonen *et al.* (2012) anticipate that where students feel socially connected and supported in school, more positive predispositions are promoted in school and classrooms in general and student motivation towards academic activities is promoted.

Reeves' (1994) spectrum of cooperative learning spans from unsupported to integral. The OAs were developed to support cooperative and collaborative learning however what actually occurs was dependent on both the teacher's implementation and the interrelationships among students. The continuum used instead therefore ranged from independent, where students work solely individually and do not confer with their peers of any level during the lesson, whether academically or socially, to collaborative. This is apparent when students discuss ideas, work together towards a common goal and support each other both academically and socially.

5.5.5 Context Novelty

Context Novelty was the single dimension that was exclusive to this study and not adopted from an existing framework. Context Novelty describes elements of an OA that are unexpected and novel to students. Though there are myriad OAs available for students and teachers to apply in lessons, many are much of a muchness and do not necessarily apply a novel context to the everyday curriculum.

Though the content coincides with the familiar physics curriculum, the context was anticipated to be less familiar or expected.

The researcher adopts the opinion of Hidi (1990) whereby presenting students with a novel lesson activity that they have never experienced and may not even expect to experience in the classroom, this can act as a trigger point for situational interest. A central characteristic of the OAs in this study is the astronomy context they apply to the physics curriculum in an effort to enhance students' learning experiences and stimulate their interest in the physics lesson. It is therefore crucial to assess the effectiveness of this context and the responses it triggers from students.

As this dimension was not present in any existing frameworks, a suitable continuum needed to be conceptualised. Given that the purpose of a novel context was to trigger students' interest, the continuum appropriately spanned from boring to interesting. Where students showed little or no interest in the context and did not appear to be curious about the topic, this would signify boring, however where students appeared interested in the topic, showed enjoyment and were curious about it, they were be considered to find the context interesting.

5.5.6 Teacher Role

Interestingly, the Teacher Role dimension was only evident in Reeves' (1994) framework. It is believed that other frameworks did not include a teacher-related dimension as they focused on an OA that to an extent, made the teacher redundant other than for behavioural management and assisting with computer and technical aspects. However, here the researcher still considered the teacher's role to be integral, as everything that takes place in a lesson is based on the decisions of the teacher. Part of the focus of this study was therefore to investigate the teachers' influence on the OA and how implementation varied from classroom to classroom. It is for this reason that this is the final dimension within the evaluation framework. Though the framework is used to evaluate the educational OAs in this study, many of the dimensions will be highly influenced by the teacher's implementation method. For example, even if an OA offers the opportunity for cooperative/collaborative learning among students, the teacher still has authority to neglect this opportunity or maintain that students work individually.

Reeves (1994) describes the teacher's role to vary from didactic to facilitative which is again largely influenced by their pedagogical approach and is interchangeable with the student autonomy dimension. A constructivist teacher should be facilitative whereas those who follow a more traditional approach will behave more didactically. Typically where student learning is autonomous, the teacher's role is facilitative, they guide and assist students in their learning but they do not entirely govern it and the classroom is largely student-centred. Where the teacher's role is didactic, it is they who are seen as the centre of the classroom. This approach is used to describe scenarios where teacher

dictate a series of facts to students for them to remember and students have no control over their learning.

However, the Teacher Role dimension is a two-sided coin. The continuum of didactic to facilitative is not only used to describe teachers' pedagogical approaches but also their relationships with their students. A didactic teacher is therefore considered not only to be didactic in their teaching but is also considered to be detached from their students, denoting the absence of a more personal relationship between them. A facilitative teacher on the other hand, not only places students at the centre of learning but also knows them on a personal level and individually. Where a teacher has built personal relationships with their students and has an awareness of their individual needs and capabilities, it is anticipated that this will also influence the Differentiation dimension.

5.5.7 Determining Overall Effectiveness

This chapter has set out six key dimensions that are employed in the construction of a framework that was used to evaluate students' learning experiences with a specific set of educational OAs. A continuum spanning from one extreme to another has been described for each dimension, this is a technique that is adopted from Reeves (1994) and used as a means of providing a visual representation of the overall effectiveness of the OAs in terms of each individual dimension. A summarising diagram of these continuums is therefore provided not only for each OA overall (see Chapters 10 and 11). The integral role of the teacher was highlighted in the preceding chapter and therefore it is expected that teacher implementation will cause positions on the continuums to vary from classroom to classroom for each given OA.

5.6 Summary

This chapter has justified the necessity for an evaluation framework in order to refine the substantial amount of data that is collected through mixed methods and through an inductive, goal-free approach. In assembling such a framework that guides the methodology and analysis of the study, this chapter has discussed the relevant theories that underpin our current understanding of students' learning experiences. The final framework is composed of seven dimensions, six of which have been adopted from existing frameworks within the literature and one that is entirely unique to this study. These are:

Goal Orientation Differentiation Student Autonomy Collaborative/Cooperative Learning Context Novelty Teacher Role Overall Effectiveness These dimensions were chosen based on their relevance and applicability to the study's research questions and are considered to be key to understanding what makes an effective OA according to the defined criteria (see Chapter 4) and that promotes positive learning experiences for students and ease of implementation for teachers.

Chapter 6

Methodological Design

This chapter argues and justifies the decision for this research study to follow a multiple-case study design. This takes the form of a pre- and post-test, mixed method approach with a sequential explanatory design (Creswell, 2003; Johnson and Onwuegbuzie, 2004; Teddlie and Tashakkori, 2009; Cohen *et al.*, 2011).

Here, the nature of the research is set out in terms of its epistemological and ontological foundations, and a detailed description of the research design and procedures for participant recruitment is provided.

6.1 The Researcher: Predispositions, Beliefs and Biases

This chapter begins by setting out the predispositions of the researcher, including her educational experiences as a student herself. This is done so that the reader can understand the motivations that led the researcher to pursue this particular research study and her beliefs in terms of how education should be delivered and received.

The necessity to clarify such dispositions becomes clear as this chapter progresses into descriptions of the nature of this study and why such approaches were taken.

Prior to pursuing this research study the researcher's education was primarily scientific, obtaining a BSc in Astronomy in 2014. However, in the early years of secondary education, science was not a subject the researcher regarded with particular interest and was a fairly average student in terms of performance. When it came to choosing GCSE subjects, it crossed her mind to pursue triple science, but there were other subjects she was more competent in and no one had really expressed how beneficial the triple science course might be. As a result she pursued additional science and it was not until the second half of year 10 that the she started to gain an interest in science and doubted her decisions against triple science. One of the physics lessons touched upon areas of astronomy, and her attention was drawn. At this point, the researcher began to focus a lot more in all aspects in the subject. The more she applied herself to physics, the more she realised she found it interesting and actually understood it, even in areas where she had less interest, she continued to work hard. This eventually transferred to her other science classes, chemistry and biology.

Reflecting on theories of individual and situational interest, the researcher considers that her own experiences could be reflective of Hidi and Rinninger's (2006) four-phase model of interest, whereby a trigger of situational interest eventually developed into a sustained, individual interest. In conjunction with a change in interest, the researcher's grades improved significantly and at the point of the first set of GCSE exams sat in the summer of year 10, she achieved an A grade. By the end of year 11 when the second half of the exams were completed, she progressed to an A*.

Despite this performance improvement and noticeably improved attention, none of it was acknowledged by her teachers. When it came to selecting A-level subjects, again, no advice or suggestions were made. However, now that certain areas of science had sparked her interest she made the decision to carry physics and biology (among other subjects) into A-level. Though the difficulty of the subjects increased greatly, she was not deterred from her fascination in astronomy. Eventually after completing her A-levels, she went on to higher education and graduated with a BSc in Astronomy with First Class Honours.

After graduation, the researcher looked into carrying out some volunteer work in science outreach. Registering as a STEM ambassador, this allowed her to volunteer at a variety of science events for various audiences such as schools and the general public. She volunteered with the Oxford Museums and also at a local library for a summer reading challenge for children. The researcher found this experience of working with a variety of audiences and seeing all the different mediums of learning, particularly interesting.

When working at science festivals, she discovered that whilst communicating general aspects of science, she was often able to apply it to contexts of astronomy and space. Not only was this her area of expertise, it seemed to trigger interest in people she spoke to. When an opportunity was presented to develop a series of astronomy-based educational materials and carry out some research, it naturally seemed like the perfect combination.

It was not until around this point that the researcher began to reflect on her previous experiences in education that she realised how little guidance she had received from her school and teachers, both when she was an average student and when she was a high performing one. This became a great frustration and she believed that everyone should be encouraged to pursue a subject if they have an interest in it.

In hindsight, the researcher's pursuit of physics and astronomy was entirely self-driven and had it not been that something sparked her interest, it would not have carried her through her particular educational path. The researcher therefore believes that greater attempts should be made to foster interest in physics and a wider variety of contexts should be applied to the physics syllabus content in efforts to do so. When transitioning into the PhD, the changeover from physical science to social science was very different and an almost entirely unfamiliar concept. Given the experience of the researcher, it would perhaps be expected that the researcher would follow a positivistic stance in the educational research, due to its similarities to research in the physical sciences. However, the opposite was desired.

Coming from a background of looking at data and observations of galaxies, the research was primarily black and white and largely objective. When the focus of the researcher turned to interactions and engagement among human participants, the researcher felt that such separate fields should not be studied with equivalent approaches. In order to successfully cross into this new research field, it was decided that the approach should also change. To fully recognise the difference in the nature of the two subject matters, it was felt that the tendencies and dispositions of the researcher when approaching research must also be altered accordingly.

Coming into the study the researcher had a clear lack of professional experience in this research area. Though she had experience of delivering workshops at CPD and teacher training events, perhaps the richest experiences were of being a student herself and learning science at secondary school. It was merely these experiences that the researcher held when designing the OAs used for implementation in this study. The researcher therefore began this study with a very different perception of how implementation may take place. For example, activities were much more involved, requiring even greater input from students and would have needed more than one lesson. Given that the majority of her experiences of education were that of being a student herself, her perspective was weighted towards student experiences. However, upon immersing herself in the classroom environment and spending time with teachers; going through their timetables, it became increasingly clear just how pushed teachers are for time and how each lesson is crucial and students must walk away from it with something that they will use in their exams. This led her perspective to shift towards the teachers and the researcher began adjusting OAs to fit within a one-hour lesson and to condense as many learning opportunities as possible into this time frame. Though teachers were still informed of all the directions the OAs could take, they were typically suggested as extension activities and evidently in no groups were they carried out in allocated lesson time.

Following an approach that placed her almost as a participant within the research environment, and subject to the same experiences as both the students and teachers, provided the best opportunity to develop this limited knowledge and understanding of educational research, pedagogy and classroom practice. The researcher no longer positioned herself as an outside observer who controls the variables of the research environment, but was instead immersed within the environment of almost infinite variables which are experienced freely.

6.2 The Nature of Research

Having set out the experiences, beliefs and predispositions of the researcher, we are brought to the nature of the study and the underpinning philosophical assumptions (Cohen *et al.*, 2011). Such assumptions stem from the epistemological and ontological predispositions of the researcher and are critical before beginning the research process.

Epistemology describes the process of gaining knowledge, its nature and how it is validated (Gall *et al.*, 1996). One who considers knowledge to be objective and is akin to procedures of the natural and physical sciences is considered to be positivistic in their research stance. The purpose of positivist research is to either prove or disprove a given hypothesis, with an emphasis on the scientific method, statistical analysis and generalizable findings (Mack, 2010). Positivists are therefore much more regimented, and often follow purely quantitative methods (Cohen *et al.*, 2011). Results are typically expressed in terms of global laws and generalisations.

One who perceives knowledge to be a subjective, personal endeavour and who insists on an involvement with their subjects is considered a constructivist or interpretivist. They argue that a researcher can only understand an individual's behaviour upon sharing their frame of reference and that the social world can only be understood by an individual who is part of the ongoing environment (Cohen et al., 2011; Burrell and Morgan, 1979). Gall et al. (1996) make the point that 'reality' is constructed differently by different individuals, therefore every individual lives a slightly different reality based on their experiences, predispositions and internal beliefs. A researcher experiences the subject phenomena from within the environment, through direct experience rather than as an external, detached observer (Mack, 2010) and theories are consequently built from the bottom, up. The researcher gathers individual perspectives from participants and from these, identifies broad patterns which eventually generate broad understandings (Creswell and Plano-Clark, 2011). This is given the broader term of phenomenology, in which one sees behaviour as determined by the phenomena of experience rather than by external, objective and physically described reality (English and English, 1958). In Chapter 3, the researcher set out a specific definition of constructivism which is the instructional design followed by the OAs of this study. This same definition is carried into the research methodology whereby the researcher's understanding of the research phenomenon and environment, occurs as a result of her experiences within it. Knowledge and understanding that addresses the specific research questions of this study are yielded by an experiential and reflective process.

Galda (2017) emphasises that where positivistic research seeks to eliminate potential biases, constructivist research, or more generally, qualitative research, does not desire to do so. The researcher is integral to the research process and therefore emphasis should instead be placed on their transparency and reflexivity. The constructivist approach therefore emphasises the importance of

disclosing all personal beliefs, values and predispositions that were set out in Chapter 6.1 and shapes the research inquiry and interpretations drawn (Creswell and Miller, 2000).

The researcher also applies inductive reasoning to such interpretations. This is an approach whereby the study presents a broad area of investigation and concepts and themes are derived from the researcher's interpretations of raw data, rather than from a pre-defined hypothesis (Thomas, 2006). Given that this study did not seek to test an existing theory but instead develop one around students' learning experiences, a deductive approach is not suitable (Elo and Kyngäs, 2008). Inductive approaches are much more open-ended and thus well-suited to exploratory research (Trochim, 2006).

The researcher therefore narrates and describes the research process as it happens and as a theory emerges it is revisited and reformed as more and more data comes to light and experiences of the researcher grows, thus influencing interpretations through reflexivity.

This "goal-free" approach (Scriven, 1991, p.56) also ensures that actual effects are measured and not just those intended or expected. As a result, less data is likely to be ignored than it would be for goaldirected, deductive approaches. The researcher considers this to be particularly important and believes that new theories cannot develop without these unanticipated effects and the presence of new, unpredicted data. This is believed to be particularly important in this study where the researcher has limited knowledge and experience in the research field. A goal-free approach therefore permits more thorough reflexivity.

Ontology describes one's view of reality and being (Mack, 2010) and is therefore concerned with the nature of the social phenomena which is under investigation (Cohen *et al.*, 2011; Burrell and Morgan, 1979). Although separate, an individual's ontological position ties closely with their epistemological beliefs, here the researcher considers a subjective reality, believing in multiple realities which are dependent on the perspectives of participants and observers (Creswell and Plano-Clark, 2011). This is contradictory to a positivist's belief in a single, objective reality.

Hitchcock and Hughes (1995) maintain that the human world cannot ever be a world in itself. Instead, it is a world that is experienced and related to a conscious subject. Data is incapable of speaking for itself and instead it comes from the researcher's interpretation who cannot possibly observe a phenomenon without altering it in one way or another (cf. quantum mechanics). Lincoln and Guba (1985, p. 37) eloquently describe this, "*The inquirer and the "object" of inquiry interact to influence one another; knower and known are inseparable.* " Words and numbers are simply symbolic representations of the real-world (Hitchcock and Hughes, 1995). Knowledge is not gained through an objective ontology and epistemology but instead through a living experience of reality. The researcher thereby agrees with the interpretivist argument that an objective description of reality is not possible and that the observer's descriptions are tinted by their individual understanding of reality in historical, cultural, ideological, gender-based and linguistic terms (Sandberg, 2005).

Particularly in the context of an educational study, where by its very nature, presents an infinite number of variables and perspectives that could all point to different assumptions and conclusions. Here, educational research is considered to be bound by time and context and subsequently more suited to constructivism as it contests the positivistic notion of generalisability. No two classrooms are identical and a class of 30 students and one teacher represents 31 different perspectives of reality, all of which are equally valid. These perspectives will vary according to, inter-alia, students' backgrounds, their lives outside of school and their social class. The teachers will vary in their years of experience, their education, their pedagogical predispositions etc. Schools also have their own characteristics and priorities such as exam results, student satisfaction and pedagogical approaches.

Having said this, it is still important that the research can contribute to field of educational research and offer relevance beyond the confinements of this study alone. The researcher therefore focuses on an alternative definition, transferability. Transferability places a much greater emphasis on the contexts and circumstances of research and is a common substitute for generalisability made by constructivist researchers.

For such reasons discussed, it was considered that a constructivist approach that applies inductive analysis to identify themes among raw data and build towards a new theory was the most suited to the researcher and the context of this study. Although it could be argued that a positivist approach could lead to more defined answers, by its nature it was considered to be too restrictive for a goal-free research scenario with no pre-defined hypothesis and that investigates such a diverse research scenario that is presented by the uniqueness of each individual environment and participant within this study (Cohen *et al.*, 2011; Burrell and Morgan, 1979). It is also considered that as the study delves into students' experiences, this is a highly complex, multifaceted construct that in itself is subjective. An objective, positivistic approach therefore does not marry with the subject matter of the study.

In conjunction with this approach, instead of assuming the role of a tester or examiner who performs research *on* their participants (positivist), the researcher adopts the role of learner, who immerses themselves within the environment, learning *with* the participants (Ryan, 2006). Where positivism upholds that the scientist, or researcher, observes an objective reality, x led to y, the constructivist applies less of a focus on the results themselves, that is, not solely that x led to y but more on the underlying processes between x and y, specifically, why did y occur? What took place along the way and what circumstances and processes are likely to have led to y. Thus, the researcher does not solely investigate *what* happens, but also *how* and *why* it happened. Nonetheless, the research is mixed methods that utilises objective quantitative measures, thus the researcher was still able to highlight the overarching element of whether x led to y.

Due to the constructivist epistemology of the research, the subjective nature of the researcher's interpretations and the problem of generalisability in educational settings, it was important for the

researcher to be immersed in the research context and environment in order to gain a first-hand experience of events and processes that surround x leading to y. This meant that rich, detailed descriptions could be produced in the data evidence and the reader can understand the context of events and how they might transfer to different classroom environments. This process enhances the transferability of the findings as it allows the teacher to better understand the context and environmental components of implementation, thus allowing them to identify the commonalities and differences of such scenarios with their own classroom. The data therefore possesses more of a narrative than a concise list of findings. This is discussed in greater depth in Chapter 8.

However, given that constructivism is built around the experiences and interpretations of the researcher, this introduces substantial levels of bias. To address the bias of the researcher's interpretations of data, reflexivity (Cohen *et al.*, 2011; Gall *et al.*, 1996; Symon and Cassell, 2012) is applied. Given that with constructivism the researcher stands as the key lens through which data is analysed and interpreted, such processes stem from pre-existing understandings, dispositions and values. Interpretations then build and evolve gradually through increased experience and maturation of the researcher (Bruner, 1986). Gall *et al.* (1996) discuss how a single object can represent multiple social realities depending on how it is interpreted by an individual.

Alvesson and Sköldberg (2009) break down reflexivity into two components: interpretation and reflection. Interpretation acknowledges that the interpretation of the researcher is not merely an analysis of evidence and data but is influenced by additional elements such as their internal values, political positions and use of language. Reflection is central to the researcher and how they must also acknowledge themselves in terms of their research community and both their intellectual and cultural customs that will also guide the research. Reflection becomes a form of interpretation and is what makes the research reflexive (Symon and Cassell, 2012).

Appropriately with the nature of this study, the predispositions of the researcher have been highlighted before the data collection is described (Sandberg, 2005). This provides a background and baseline for the interpretations that are later made in the analysis so that the reader has a clearer understanding of the analysis and interpretation processes.

6.2.2 Summary

Given the circumstances, this study applies a constructivist epistemology that is complemented by an inductive, reflexive approach to analysis and goal-free evaluation. As highlighted, the researcher developed the educational OAs with a strong background in astronomy and science, but limited experience in education. Although previous literature has been reviewed and the researcher has previously participated in delivering astronomy-based workshops at various teacher-training events, the majority of her experiences are as a student herself. A specific hypothesis cannot be generated where little experience is held by the researcher in regards to the theory behind the process. This

places the researcher as a learner within the research environment, who discovers the underlying processes along the way and through their experiences with the participants. These experiences are thoroughly narrated and evaluated along the way in order to build a theory around them. This 'diary' method allows the reader to follow the researcher's train of thought. It is also helpful in limiting the danger of the researcher distorting evidence and data due to unconscious biases (Verma and Mallick, 1999). Researchers will typically hold some hope that the findings will point in a particular direction. For such an in-depth, personal approach, with data collection methods that are highly qualitative, there is a risk that responses could get misinterpreted, due to this desire to find a particular result.

However, despite the largely open-ended constructivist approach of this study, it still held a specific focus with particular research questions.

6.3 Research Questions

This study investigates secondary school students' perceptions of their learning experiences in physics at school and during the implementation of educational OAs developed around an astronomy context and an inquiry-based pedagogy. Due to the nature of this study and the epistemological assumptions of the researcher, research questions posed by this study do not set out to simply ask '*do*' and '*what*' questions but instead ask the '*how*' and '*why*' behind the specific outcomes that come to light.

The study has four key research questions:

- 1. How are students' perceptions of positive learning experiences promoted within a physics lesson?
- 2. How does teachers' planning and delivery of the Down2Earth and Hubble outreach activities influence students' learning experiences?
- 3. How does the implementation of the Down2Earth and Hubble outreach activities reveal the opportunities and limitations they lend to students' learning experiences?
- 4. What are the implications for the future development and utilisation of physics outreach activities for developers and teachers?

This chapter will detail the methods that are considered to be most suitable for answering these four questions in line with the nature of this study.

6.4 Research Design

The defined nature of the research is used to guide the research model that is followed throughout this study. Research in the social sciences has developed and progressed significantly since its foundation,

and various research models have evolved. These models create a foundation of concepts and techniques that guide a research study (Hitchcock and Hughes, 1995). There remains an ongoing debate in the literature around which model is the most appropriate and under what circumstances.

This research study will follow a pre- and post-test mixed method approach through a sequential explanatory, multiple-case study design (Johnson and Onwuegbuzie, 2004; Creswell, 2009; Teddlie and Tashakkori, 2009; Cohen *et al.*, 2011). The following subchapters break down the three components of this study's design and argues the suitability of each.

6.4.1 Mixed Method Approach

In terms of the mixed method element, this is an approach that has emerged over the last 20 years (Teddlie and Tashakkori, 2009) and combines quantitative and qualitative research paradigms into one study (Creswell, and Clark, 2011; Johnson and Onwuegbuzie, 2004; Teddlie and Tashakkori, 2009; Cohen *et al.*, 2011). Johnson and Onwuegbuzie (2004, p. 14) describe the approach as "*the third research paradigm in educational research*".

Muijs (2004) describes quantitative approaches as processes that explain a particular phenomenon through numerical data collection such as tests and closed-questionnaires. Quantitative research is seen as an objective approach to social phenomenon (Firestone, 1987; Teddlie and Tashakkori, 2009) and is well suited to studies that test hypotheses and measure numerical change (Muijs, 2004). It is therefore recognised as the reciprocal of positivistic research. Although the researcher considers quantitative approaches to be valuable in identifying the 'big picture' and useful in informing other inquiries, entirely quantitative approaches are not considered to be compatible with the nature of this study or capable of fully addressing the necessary '*how*' and '*why*' aspects.

Alternatively, qualitative methods collect non-numerical data, typically through interviews, observations and documentation (Denzin, 1978; Punch, 2009), and take a more subjective stance to research. Interviews provide a medium to access individuals' perceptions, understandings and constructions of reality (Punch, 2009). Jones (1985) explains that in order to understand something, participants must be able to express their views in their own way, rather than through a restricted and closed format such as that offered by a questionnaire. With particular regard to educational research, Erickson (2012) underlines the necessity for qualitative methods when looking for detailed information and to identify and understand changes with time.

The third mixed method paradigm first originated from Campbell and Fiske in 1959 who applied a *"multimethod matrix"* to investigate the validity of psychological traits (Campbell and Fiske, 1959). They argued that triangulation of mixed methods in validation processes strengthens validity and yields results that are not possible from singular methodological procedures (Johnson *et al.*, 2007).

Denscombe (2008, p. 272) highlights various qualities that are achieved through a mixed methods research design. Those that are of particular relevance to this study are:

1/. To improve the accuracy of data

2/. To provide a more complete picture of the subject through combination of different sources and methods

3/. A way of compensating for and overcoming limitations of single method approaches

4/. A way of developing analysis procedures and expanding on findings from initial phases

This third paradigm reflects a world that is neither entirely quantitative nor entirely qualitative (Cohen *et al.*, 2011). Instead it is a combined and complimentary approach that presents the best prospect for addressing research questions and gathering useful answers (Johnson and Onwuegbuzie, 2004; Ivankova *et al.*, 2006).

It is this third mixed methods research paradigm that was considered to be most applicable to this particular study. There are two key points in the rationale of this decision. For one it offers an avenue to combine two separate paradigms that can surmount the limitations of single method approaches. As a result, we are provided with a more holistic understanding of the subject phenomenon (Johnson and Onwuegbuzie, 2004; Ivankova *et al.*, 2006; Denscombe, 2008; Greene, 2008). Secondly, Chapter 6.1 emphasises that where the researcher was transitioning from the field of astronomy into that of social science she considered that shifting her approach to research could enable a smoother, more informed transition. Accordingly, where the researcher is immersed within the research environment and obtaining mixed sources of evidence, it is considered to be advantageous in developing her understanding of school settings and enriching her interpretations through reflexivity of her experiences.

6.4.2 Sequential Explanatory Design

The use of a mixed method approach also allows for a sequential explanatory design. This is a term coined by Creswell and Plano-Clark (2011) who describe a series of sequential and concurrent strategies for carrying out mixed method research. Sequential strategies involve multiple phases of data collection, beginning with one method (either quantitative or qualitative) and then following with another at a later time.

Sequential methods can be either explanatory or exploratory, depending on the sequence of methods (Creswell and Plano-Clark, 2011). Where exploratory methods implement qualitative methods first and follow these with quantitative methods in order to seek a generalisation of their initial results, an explanatory method implements the quantitative phase first and then applies a qualitative phase in order to understand the patterns in the data and the reasons behind such findings. A sequential

explanatory design is therefore much more reflective of the nature of this study that set out to implement quantitative questionnaires in order to assess the 'big picture' of students' learning experiences. This was then complemented with qualitative classroom observations and followed up with focus groups that were informed by the quantitative data.

Once all the data had been collected, quantitative and qualitative data was initially analysed separately, and a between-methods triangulation of this data was carried out (Denzin, 1978) that fed into the evaluation framework described in Chapter 5. The researcher (as with others) is of the belief that between-methods triangulation is a stronger approach when compared to within-methods triangulation (multiple uses of a single qualitative or quantitative method), as it is not limited by its paradigmatic approach. Often it is found that the deficiencies of one method are the strengths of the other (Denzin, 1978; Johnson *et al.*, 2007).

6.4.3 Case Studies

According to the English Oxford Dictionary (2018), a case study is defined as:

A process or record of research into the development of a particular person, group or situation over a period of time

Stake (1995, p. xi) describes a case study as a study of the "*particularity and complexity*" of a given case and the search for detail and understanding of its activity in particular circumstances and interaction within particular contexts. Case studies are therefore highly in-depth procedures and what they lack in breadth they make up for in depth. Yin (2009) considers case studies to be well suited to a mixed method approach as they have a unique strength that allows them to deal with a variety of evidence and formats of data (Yin, 2009), therefore providing a strong foundation to gather data of both quantitative and qualitative strands.

Stake (1995) posits that case studies, although guided by specific research questions, do not benefit from a rigid structure and should be flexible in order for the researcher to make changes throughout the process. He justifies this stance by drawing on notions from Parlett and Hamilton (1972) who argue that the course of which a study takes cannot be known beforehand (Yazan, 2015). This idea of a flexible design was also believed to be compatible with inductive reasoning and reflexivity that was implemented in this study. Given the initial limited experience of the researcher in educational research, her understandings and interpretations of school environments evolved rapidly as her exposure to such settings dramatically increased. This in turn informed the future direction of the study and allowed for the gradual development of a theory.

Although case studies are well established in qualitative educational research, there is no unanimous agreement on how they should be designed and implemented (Yazan, 2015). Different dispositions are again largely based on the epistemological beliefs of the researcher. Two esteemed case study

methodologists are Yin and Stake, where Yin employs a predominantly positivistic view, Stake is a constructivist. It is therefore Stake's propositions that were deemed the most applicable and complementary to the nature of this study. Though despite various epistemologies, advocates largely agree on the advantages of case study research. Their ability to deal with questions that address operational events that occur over time, qualify case studies in education a much deeper, more personal perception of both teaching and learning, thus helping towards unravelling the opinions and behaviours of both teachers and students (Verma and Mallick, 1999).

However, as with most approaches, case studies are not without their limitations and due to their small, isolated samples, they are most commonly criticised for their inability to generalise. However, as touched upon previously, this study seeks to offer transferability rather than generalisability and thus requires the researcher to acknowledge contextual factors within the research environment. A case study design is well-suited to such a requirement as it allows the researcher to interrogate the scenario more deeply, thus achieving a richer understanding of underlying constructs of students' learning experiences than could be achieved in a broader, less intensive approach.

Verma and Mallick (1999) maintain that the utmost advantage of a case study design is its venture to understand an individual as a whole, with relation to his or her environment. In this context, the individual OAs (Down2Earth and Hubble's Expansion of the Universe) stand as the cases, and the aim is to understand students' learning experiences in terms of their everyday physics lessons and how they respond to the educational material provided in implementation.

This study therefore addresses '*how*' the educational materials are implemented and '*how*' and '*why*' they induce particular learning experiences among the students.

Cohen *et al.* (2011, p. 54) emphasise that educationalists are not merely interested in '*what works*', but also '*why*' and '*how*' it works, for *whom* and under what given circumstances. This is summarised in the nature of the research questions this study sets out. However, a common misconception of the case study design is that it is simply a description of an individual, event or a situation; and could also be argued the case of Scriven's goal-free evaluation. However, this is countered by Verma and Mallick (1999) who emphasise that the focus of a case study is instead on the interactions between these three factors. Hitchcock and Hughes (1995) also stress that case studies concentrate on understanding how particular situations and events converge to create particular outcomes. This was particularly applicable to this study's framework dimensions many of which were closely related.

These key features position the case study as a strong design format for this educational study. However, these arguments from Verma and Mallick (1999) and Hitchcock and Hughes (1995) are not a given, but was the responsibility of the researcher to ensure the case studies were not mere descriptions, but clearly laid out the relationships between the individuals, the event and the scenario. In terms of the two cases in this study, as investigation focused around the implementation of different OAs, a clear separation had to be defined. The Down2Earth and Hubble OAs are very different, are designed for different audiences and apply to different areas of the physics curriculum. They therefore cannot be analysed or evaluated collectively and must instead be separated.

Therefore, each case (OA) is implemented across a number of groups, (typically classes from different schools). Both quantitative and qualitative data collection methods were employed for each case and for each group, permitting the investigation of more complex scenarios. In turn, this expanded both the breadth and scope of the case study (Townsend, 2015).

6.5 Recruiting Schools, Teachers and Students

Overall, this study recruited 11 groups of participants. Of these groups, nine were a standard school class and of the other two, one comprised of students from the same school across different year groups and the second were students from different schools but were in the same year group. This study primarily (though not exclusively) employed opportunity sampling in order to recruit participants. This method involves recruiting people from the target population that are willing to take part. This method of sampling is common in educational research as it is one of the more convenient methods. As this is a small-scale initiative, teachers are the gatekeepers to the students and classroom and therefore need to be willing to participate and dedicate their lesson time to the study and implementation of the OAs.

It is recognised that this method presents several limitations. Participants are of a biased sample that is not representative of the entire population. For example it could be that only teachers from high-performing schools would be comfortable in taking part and committing their lesson time to an alternative teaching method. It could be that such teachers are already applying measures to ensure that they are delivering innovative lessons and actively seek new OAs and alternative teaching approaches and mediums. Alternatively, lower-performing schools that are faced with a larger pressure on accountability and performance are likely to be more resistant. However, the demographics of the participating groups are described in Chapter 6.7 and it was seen that this was not necessarily the case. What's more, and reducing the bias of the participants in this study, is that the Educational Achievement Service (EAS) in Wales, (an organisation intended to improve the standards of education across South East Wales), took an interest in the study and recruited several schools on the researcher's behalf. These school were therefore not necessarily entirely willing but are merely encouraged to follow the guidance of the EAS.

Elements of purposeful sampling were also used. This is an approach that is typically applied when the researcher seeks to select scenarios that are "*information-rich*" (Gall *et al.*, 1996, p. 218). It is

therefore more typical for small-scale qualitative studies, where the purpose of data collection is to foster a greater understanding of the subject phenomena. A broader range in participant groups (schools or classes) also enhances the transferability of research findings and increase the likelihood of teachers identifying scenarios within the study that are applicable to their own classroom. In this instance, schools with varying characteristics were sought. This included a range of free-school-meal eligibility, a combination of public and state schools and also a combination of single-sex and co-ed schools. Unfortunately the researcher was only successful in recruiting all-male schools and no all-female schools participated in this study.

It was discovered early on that getting into schools was extremely challenging and so a number of methods and mediums were used in order to increase likelihood of recruiting groups, such as that of getting the EAS on board with the study, direct emails to schools or teachers, physics teacher conferences, TeachMeets, Cardiff University blog, word-of-mouth and other additional access routes. Each of these is described in Appendix 3 and the invitation email sent to teachers is provided in Appendix 4.

6.6 Ethical Procedures

However, before any recruitment could begin, it was essential to obtain ethical clearance for the research study. Ethical procedures, particularly where human participants are involved are crucial. The British Educational Research Association [BERA] (2018, p.4) sets out five key principles that should be followed in all social science research:

- 1. Research should be inclusive of different interests, values, funders, methods and perspectives in order to align with a democratic society
- 2. Respect the privacy, autonomy, diversity, values and dignity of individuals, groups and communities
- 3. Research should be conducted with integrity throughout, employing the most appropriate methods for the specific research purpose
- 4. All social scientists should act with regard to their social responsibilities in conducting and disseminating their research
- 5. All research should aim to maximise benefit and minimise harm

Alderson (2005) emphasises that ethical considerations should not be a final hurdle in research design but are a vital component that set the standards of the study. Ethical processes force the researcher to consider two crucial questions, is the research worth doing? And is the research explained clearly enough so that invited participants can make an informed decision around their consent? Before implementation of this study, a series of ethical procedures were therefore carried out in order to confirm that the research was respectful of: the people involved, democratic values, the quality of research, and academic freedom (BERA, 2018).

Although the main participants of this study were students, teachers were the primary avenue for recruitment. Teachers were approached through a number of methods, all of which are listed in Appendix 3. However, under all conditions where teachers were invited to take part in the study, it was made clear that participation was not compulsory in order to use the OAs.

Under circumstances where teachers expressed an interest in participation, they were then formally invited to partake in the study. Information sheets were provided to teacher and students, and where students were under the age of 16, to parents. This included a formal invitation and detailed information regarding the following aspect of the research process (wording was adjusted accordingly for teachers, students and parents/guardians):

What are the astronomy and physics educational materials?
What is the research?
Who is being invited to participate and what will it involve?
Are there any risks or benefits to me or my students from taking part in this study?
What happens to the information me and my students provide?
What if I or my students don't want to take part?
What if I change my mind about participation?
Confidentiality
Contact Information

Participants were informed on exactly what would be asked of them, how it would affect them and what to do if they did, or did not, want to be involved. They were also reminded of ensured anonymity, their right to withdraw at any time with no consequences and were provided with the researcher's contact information should they have any queries. Attached to each of the information sheets was a consent form.

The consent process of the study involved an opt-in procedure for teachers on behalf of themselves and their students, and an opt-out process for students or parents/guardians. This meant that teachers provided consent on behalf of themselves and their students but should a student, or a parent of a student under the age of 16, not want to participate, they would opt-out of the study by returning the provided consent form. Should no forms be returned from students or a parent/guardian, consent was considered to be granted by their teacher. This method was chosen based on discussions with teachers, where it was evident that this was the most preferred approach and their typical protocol. All data that was collected was completely anonymous and non-identifiable. Where focus groups or interviews were carried out, they were recorded and anonymised on transcription. This study was granted approval by the ethics committee in the School of Physics and Astronomy at Cardiff University. The ethical form application and approval documents are provided in Appendix 5 and 6. Full information sheets with accompanying consent forms can be found in Appendix 7.

All data reported in this thesis is entirely anonymous and not identifiable to anyone personally. Each school who participated in the study was assigned a colour, this colour was also assigned as a pseudonym to the name of the teachers. For example one school was group Green and their teacher was Mr Green. Where specific students are mentioned they are also given pseudonyms. This process was done in order to protect the identity of schools, teachers and students. Although the schools were not named specifically, they are described in Chapters 6.7 and 6.8 in terms of a number of characteristics, such as their approximate location, size and free-school-meal (FSM) eligibility, this is done in line with the constructivist nature of this research and to 'set the scene' for the reader so they have an understanding on the research environment and school settings.

6.7 Case Study 1: Down2Earth

Five groups (school classes) were recruited in total for the Down2Earth case study. Their key characteristics are summarised in Table 7 and the following subchapters provide a more detailed context for each group. For points of reference, the national rate for FSM eligibility among students across all compulsory school years is 17.0% in Wales (Welsh Government, 2017b) and 29.1% in England (GOV.UK, 2017). The average pupil teacher ratio is 16.5 in Wales, a value for the England equivalent could not be found.

6.7.1 Group Pink

Group Pink was a class from an all-boys public school in England with approximately 1,000 students. The school was visited for implementation in July 2017.

The teacher that served as the point of contact at school Pink has been involved with the Faulkes Telescope project, on-and-off for several years, though this was mostly before the researcher's involvement in the project. However he was interested to hear about more recent developments and arranged a meeting with the researcher at the school to discuss the study.

The teacher was happy to incorporate the Down2Earth activities into a lesson with a group of year 9 students and even arranged to change their usual lesson timetable as he felt the students would benefit from a double lesson (two hours) which was agreed with another teacher who had the class before him. As both teachers were only available for their usual lesson, the researcher delivered the activity

to the students and the second, female teacher was present for the first half, and this male teacher (the point of contact) was present for the second.

This class completed both the Making a Meteorite and Recreate a Crater activities and also had some time to do some extension activities.

6.7.2 Group Green

Group Green was organised by the teacher from group Pink who often works with this local state school and had arranged for the researcher to visit the school immediately after case Pink (July, 2017). Group Green was a year 9 class from a co-ed school with a total of 781 students in the 2016/17 academic year. Their most recent Ofsted inspection was in 2015 where they received a 'Good'. The overall FSM eligibility of the school is 22.1%, below the national percentage of 29.1%.

Group Green had a female teacher who had had no contact with the researcher and all arrangements had been made through the teacher from group Pink, therefore the researcher was asked to deliver the activity to students.

Though implementation was scheduled for regular physics lesson period, students were in a computer room rather than their usual classroom. The lesson was an hour long so the researcher planned for the students to complete the Making a Meteorite activity. Some students progressed a lot quicker than others and did move on to the Recreate a Crater activity as well.

6.7.3 Group Blue

Located around the Newport area in Wales is school Blue. The co-ed school hosts 876 students and has one of the highest FSM eligibility at 30.9%. The pupil teacher ratio is 15.4 with 57 teachers, this ratio has declined since 2010 when it was 18.5. At present, this ratio is below both the Wales average and the local authority average (16.9).

The teacher at this school was put in contact with the researcher through the EAS Science Advisor for Newport who joined the researcher in visiting the school to meet with the teacher.

The participating class in this study was a top set year 9 science class of 31 students, 15 male, 16 female. The class had a male teacher and also a male teaching assistant/support staff. The teacher had gone to extra lengths to incorporate this activity into the school timetable as he wanted the students to complete the activity over two lessons that were close together. He therefore arranged with the class's chemistry teacher to have two consecutive physics lessons, rather than physics followed by chemistry.

Unfortunately, more rearrangements had to be made with the second lesson due to examinations and the second lesson ended up being 2 weeks later.

The first lesson was implemented as an introductory lesson to the subject matter. The teacher introduced students to some topics of astronomy and the researcher demonstrated the Impact Calculator to the students, explaining how it works and how they would be using this 'virtual lab' to carry out an investigation for the Making a Meteorite activity that involves investigating what happens to the kinetic energy and size of the impact crater when you alter various parameters.

The teacher split the class into four groups and told each group what parameter they would be investigating (projectile diameter, projectile velocity, trajectory angle or projectile density). He then drew a table on the board for students to copy into their books, ready to record their data into. The activity suggests that students design their table of results themselves, but here, the teacher decided that this group of students needed additional support and guidance.

Students then had to formulate a hypothesis of what they expected to happen when they changed their specific parameter.

The investigation itself was then carried out in the second lesson. Here they collected their data, recorded their results, drew conclusions and compared them to what they had originally predicted, though not all students got this far.

6.7.4 Group Navy

Group Navy was the first school to be visited in this study and the first to use the educational OAs. The teacher here had attended a CPD event with FTP in the past and so was familiar with the project.

The Wales-located school has a FSM eligibility of 21.9%, above the overall percentage in Wales and more-or-less equal with the school's local authority (21.4%). However the school is above their local authority (17.1) and Wales (16.5) for their pupil teacher ratio at 18.8.

The male teacher and a group of 26 year 9 students participated in this study. Upon meeting with the teacher in the school, the teacher chose to implement several of the Down2Earth activities across multiple lessons, though the researcher only observed one. This was the Recreate a Crater activity that was implemented in November, 2016.

The teacher had arranged to book the schools' Chromebooks for this lesson in order to access the online simulation which took place in students' normal physics lesson period.

6.7.5 Group Brown

Group Brown was the final group in the Down2Earth case study. The England school hosts 1,208 students, 17.3% of whom are eligible for FSM. This statistic is below the England national percentage.

This school was recruited on a more personal level in that the researcher knew a teacher at this school who had just completed his NQT year. The researcher showed the teacher the research project she was carrying out and the educational OAs and he was keen to implement the Making a Meteorite activity with a group of year 8 students. Also of interest given the reports around the lack of physics specialists, is that this teacher was a chemistry specialist.

Though it could be considered that as the researcher knew the teacher outside of the workplace environment, it could have been a biased case, it was the students who were the primary source of data who had no previous affiliations with the researcher. The teacher was also not aware that the OAs had been developed by the researcher herself and was merely informed they had been developed by the Faulkes Telescope Project.

This was the second group whereby the researcher was able to observe the students in a normal dayto-day lesson, prior to implementation which took place over a usual physics lesson period in February 2018. Here, the teacher had booked the school laptops for students to use in order to access the online simulator.

6.8 Case Study 2: Hubble's Expansion of the Universe

The Hubble case study recruited six groups, four of which were standard school classes, one was a group of students from the same school but different year groups and one was of students from different schools but the same year group.

6.8.1 Group Yellow

School Yellow is a co-ed school located in Wales with 1,546 students and a FSM eligibility of 14.7%. There are 89 equivalent full time teachers (53 female, 36 male), yielding a pupil teacher ratio of approximately 17.4, a value that has remained relatively consistent since 2010 and exceeds both the Wales average (16.5) and the local authority average (16.1).

Participating in the study was a female physics teacher and a year 13 class comprising of one female and eight male students. Students carried out the Hubble's Expansion of the Universe activity in one of their usual lesson slots at the end of March, 2017. The teacher had previously worked with the Faulkes Telescope Project in previous years, with other members of the team (not the researcher) and so was contacted by email in order to see if she would be interested in these newly developed OAs.

The teacher was unable to get computer access for the students and so came up with a method of carrying out a paper-based version of the activity, and put the main spreadsheet onto her computer and onto the projector for the students to see.

Students were given handouts of the supernova light curve data with the light curve already produced (two supernovae each) and a copy of the table that included the radial velocity values for each supernova target's host galaxy. Students had to identify the peak magnitudes of their supernova targets and identify the radial velocity from the table. The teacher would then go around the class and students would provide their magnitude and velocity values for the teacher to input into the spreadsheet. The teacher used the spreadsheet to automatically calculate distances.

Once all the students had contributed their data into the spreadsheet, they then used these values to hand-draw their Hubble plot onto graph paper.

The teacher of Group Yellow repeated implementation with another class the following year. On the second occasion, the teacher implemented the activity slightly differently and had the students apply the distance modulus equation and call out their answers for Mrs Yellow to check on the automated spreadsheet.

6.8.2 Group Orange

School Orange is a co-ed school with 1,518 students and is located in Wales. Of all the state schools in this study (across both cases), Orange had the lowest FSM eligibility at just 6.1%, and with 81 teachers, it has a teacher ratio of 18.7. This has increased since 2010 and is way above the average for both Wales and the local authority.

A male teacher from this school responded to the researcher's post on the Cardiff University outreach blog and asked her to come in to discuss the material. The teacher was very keen to implement some of the material and said that another teacher was due to teach content from the Hubble activity the following week and that he would speak to the teacher about implementing the activity.

This female teacher agreed and implemented the activity in a year 13 physics lesson with six students. Participation took place in May 2017, when the students were covering the final areas in their syllabus before preparing for their exams.

For groups where the activity was ran by the teacher, the researcher met with them prior to implementation in order to go through the activity, or had conversations via telephone or email. However, in this instance, the teacher who met with the researcher was not the one who implemented the activity. There was therefore slightly less coherence and the teacher was evidently less prepared than in other circumstances. Despite this, she was only too happy to deliver the activity.

The students had a double period in the morning giving them two hours to do the activity and the teacher had booked laptops for them to use.

6.8.3 Group Red

Group Red came from an all-boys public school in Wales. The teacher from school Red also responded to the Cardiff University blog post about the material. Although the researcher did not meet with the teacher prior to implementation, there were several lengthy email discussions and the teacher ran through the activity himself before implementation to ensure coherence.

The teacher had two year 13 classes and wanted to implement the Hubble activity with both groups and was implemented during normal lesson periods. The activity was therefore implemented with both groups, consecutively in April 2017.

The teacher highlighted to the researcher that this was going to be implemented as more of a "revision class" as the topic had already been covered but he felt that the students would benefit from this reinforcement. The teacher therefore did not use the introductory PowerPoint with either classes and just briefly explained what they would be doing before diving straight into the activity.

6.8.4 Group Black

Group Black came from the largest school that participated in this study with a total 2,202 students. Of these, 11.8% qualify for free school meals.

This school also has the largest pupil teacher ratio out of participating schools. With a ratio of 19 they have 116 teachers (70 female, 45 male). This ratio has increased since 2010 and is now well above the Wales and local authority averages. In their most recent Estyn inspection in 2016 they received a 'Good'.

The teacher at group Black has worked with the researcher on one occasion before and with other members of the Faulkes Telescope team on several occasions. The teacher was interested in the Hubble activity and arranged for the researcher to visit the school at the beginning of May 2017.

At this stage, year 13 students had covered this area of the curriculum and so instead the teacher wanted to use this as an opportunity for almost an advanced activity for some of the more able students. He gathered a mixed group of students from year 10, 12 and 13 to partake in the activity which was delivered by the researcher.

These students were pulled out of their usual timetabled lesson to attend this and were also in the library in order to access computers, rather than in their usual physics classroom.

6.8.5 Group Purple

Group Purple was one of two groups whereby students did not all come from the same school. This was a twilight event for 'gifted and talented' year 12 students as part of the SEREN Network. The network is funded by the Welsh Government and aims to support able students in pursuing their full potential. The SEREN network is spread out across several 'hubs' throughout Wales, and the

Education Achievement Service Science Advisor put the researcher in contact with the individual who runs the South East Wales Consortia which encompasses Monmouthshire, Torfaen, Blaenau Gwent, Caerphilly and Newport. This contact wanted to implement the Hubble activity in an after school workshop. A school from one of these locations hosts the twilight event and students from other schools are able to attend.

This group therefore had a mix of students from different schools but all of high ability. All students had chosen to take at least one science at A-level but not necessarily physics. This was the largest implementation group across both case studies with approximately 60 students.

The only teacher that was present for this event was the head of sixth form at the host school, the activity was therefore ran by the researcher and given the large group, was assisted by two lecturers from Cardiff University.

This was also the first group where no computers were accessible. Given that this activity is entirely based on computer access, it to be amended in order for the activity to be carried out with just pens, paper and calculators. This proved to be a challenge to the researcher but provided an interesting and valuable comparison.

This group participated in the study in January 2018 as a two hour twilight event from 6-8pm.

6.8.6 Group Grey

Group Grey was another SEREN event but this time was in the Cardiff hub and encompassed a different network of schools. This was the smallest participating group with just five year 12 students from three different schools. The event took place at a college which none of the participants attended.

This implementation took place the day after case Purple (January 2018) and although this group had access to computers, based on the success at case Purple it was decided that this group could also do the paper version of the activity. Again the initiative activity was delivered by the researcher with an accompanying lecturer from Cardiff University. The twilight took place over two hours from 4-6pm. An observation was not carried out for this case and therefore only questionnaire data was obtained.

6.8.7 Summary

The previous sub-chapters have provided an outline for each of the groups and host schools who participated in this research study. In total, the researcher collected data from 237 students across 11 groups.

Table 7 summarises some of the key points for each group, including the location of the school, the number of students, their year group, what activity they completed and who the activity was delivered by.

Group	No. Students	Year Group	FSM Eligibility	Gender Mix	Public/State	Latest Ofsted/Estyn Rating	A D
				Down2E	Earth		
Pink	16	9	-	All Male	Public	Outstanding	Re
Green	32	9	22.1%	Co-ed	State	Good	Re
Blue	31	9	30.9%	Co-ed	State	Adequate	Т
Navy	26	9	21.9%	All Male	State	Adequate	Т
Brown	25	8	17.3%	Co-ed	State	Good	Т
			Н	ubble's Expansion	of the Universe		
Yellow	9	13	14.7%	Co-ed	State	Adequate	Т
Orange	5	13	6.1%	Co-ed	State	Adequate	Т
Red	16	13	-	All Male	Public	Excellent	Т
Black	8	10, 12, 13	11.8%	Co-ed	State	Adequate	Re
Purple	60	12	-	Co-ed	State	N/A	Re
Grey	5	12	-	Co-ed	State	N/A	Re

Table 7 – Summary of Participating Groups
6.9 Summary

This chapter has provided much detail into the methodological design of this research study. Due to the constructivist nature of the research, the experiences and viewpoints of the researcher herself have been discussed in order to provide the reader with an awareness of her predispositions and the reflexivity in the interpretations of qualitative data.

The case for using a mixed method approach has also been argued and justified in terms of how both qualitative and quantitative methods best suit the demands of the research questions of this study, enable the researcher to gain a deeper understanding of the complex scenarios in classroom environments and promote the most effective transition into a new research field.

An inductive approach is applied to data collection and analysis in order to allow theories to emerge accordingly. Methods for school and teacher recruitment have been highlighted along with an overview of each of the groups that participated in each case study. These groups have been explained as much as possible in order to provide a context for the reader without exposing the identity of these schools, their staff and students. This context is essential in order for the reader, particularly if they are a teacher, to understand the environment in which implementation took place and to relate such scenarios to their own classroom and students.

This chapter should have served towards 'setting the scene' for the reader by providing them with an in depth insight to researcher, the procedures and the environment of each group. This provides a background understanding of the nature of the researcher study before progressing to findings and results of implementation.

The following chapter details the individual instruments used in the study for data collection and justifications for such tools.

Chapter 7

Instrumentation for Data Collection

This study set out to investigate secondary school students' perceptions of their learning experiences in physics lessons and how this compares to their perceptions and experiences during the implementation of the astronomy-based educational OAs. However, it must be emphasised that this study did not set out to measure attitude change. Instead it *compared* students' perceptions and opinions towards their day-to-day physics lessons and the outreach activities they experienced. The process uncovers what features promoted positive responses and student engagement and what experiences were associated with negative experiences. This enabled the development of a deep understanding of how students want to, and perhaps should, learn physics in school.

The previous chapter set out the nature of this research study and the decision to follow a case study design with a mixed method paradigm. This chapter will now review some of the common methods of perception measurement before setting out the specific instruments that were employed in this study to gather the necessary data for answering the specific research questions.

7.1 Measuring Perception

According to the English Oxford Dictionary (2018), the term 'perception' describes "*the ability to see, hear, or become aware of something through the senses*" and "*the way in which something is regarded, understood or interpreted*". The researcher acknowledges that peoples' perceptions are closely linked with their attitude, a term which is described by the English Oxford Dictionary as "*a settled way of thinking or feeling about something*". Much of the literature reviewed in Chapter 2 focused on students' attitudes towards science and physics. However, due to the nature of the research, the constructivist epistemology, and accountability pressures faced by teachers, here the researcher has chosen to focus on students' perceptions rather than attitudes. Though related, a clear distinction is considered to lie between them. An individual's attitude is generally consistent over time (Reid, 2006) and therefore can be difficult to change (Bartlett *et al.*, 2018), this is also implied by the "settled" component in its definition. The researcher considers one's perception on the other hand to be much more context-based and shaped by the circumstances and environment at a specific point in time.

For example, even where a student may display a negative attitude towards physics at school, their perceptions will likely vary from lesson to lesson, depending on a variety of factors such as the lesson content, their mood or that of the teacher, and the format of the lesson. Though in general terms they may have a dislike towards physics, they are likely to have slightly different experiences, each time they attend a physics lesson. This also goes the other way, in that where a student may have a positive attitude towards physics, this is not to say that for the same reasons, they will never have a poor perception of a lesson. It is important to make such clarifications of these terms and what specific construct is being measured, otherwise it can lead to problems in validity and reliability of the research (Gardner, 1975; Barmby *et al.*, 2008).

Given the pressure and time restraints on schools, a longitudinal initiative is not feasible and the researcher anticipated that such a commitment it would attract less teachers. The OAs were therefore designed to span just one lesson. The researcher did not seek any attitudinal change, but instead set out to identify components of a lesson that can successfully promote positive learning experiences among students so that their perception of that given physic lesson is positive, regardless of their overarching attitude. At the most, the researcher anticipated that perhaps such experiences could encourage students to reconsider their preconceptions. If perceptions are positive, and the elements of the OAs that are seen to promote a positive learning experience are reinforced and sustained, it is possible that this could lead to a gradual change in students' attitudes in the long-term.

The researcher also considers that perceptions and experiences are arguably a more influential focus than attitudes in terms of educational practice and broader impact. In measuring students' perceptions of their learning experiences, that are context driven, there is a greater focus on such specific experiences, what specific cues led to what responses. In turn this provides smaller, more manageable techniques that teachers can adopt in their classroom.

It has been discussed that a mixed method approach to data collection that is both sequential and explanatory within a case study design is employed. One method is used to guide and enhance another method and is triangulated in order to build an evidence base and to identify confirmations and divergences in the data. This technique is particularly advantageous when addressing a latent construct, as with that of peoples' perceptions. A latent construct is not a variable that can be measured directly or addressed by just a single question. Instead it can only be inferred from other evidence from multiple indicators (Reid, 2006). This is therefore another reason for why a constructivist approach is well suited to this study.

Having emphasised the distinction between attitude and perception, it is still considered; particularly from a quantitative side that the two are measured through similar techniques. The researcher believes that the key difference lies in the post-test that is implemented subsequent to implementation. Were the researcher measuring *attitudes towards physics*, the post-test would be identical to the pre-test. For

example if a pre-test item was 'physics lessons are interesting', this same statement would be included in the post-test. Where the researcher is measuring *perceptions of learning experiences*, the focus of the post-test shifts to specific circumstances during implementation. Therefore this item becomes 'the physics activity was interesting'. The pre-test therefore measures students' perceptions of their learning experiences in day-to-day physics lessons, and the post-test measures their perceptions of the implemented OA. Though only examples, these statements are reflective of a surface level understanding of students' perceptions and is why this study employs a mixed method approach with both quantitative and qualitative methods.

Qualitative data is primarily collected through classroom observations and focus groups, although some open-ended questions are included in the student questionnaires. Observational data was collected both before (where possible) and during implementation in order to delve beyond the surface-level understanding and into the processes and experiences behind students' questionnaire responses. Qualitative methods offered a better way of understanding what specific components of their lessons or the OAs promoted or discouraged positive learning experiences and why. Post-implementation focus groups were then performed in order to follow-up any elements that came to light in questionnaires and classroom observations that were not fully understood or triggered further questions from the researcher. The qualitative side focused more on a means of interpreting the results, probing into why particular responses were or were not seen (Firestone, 1987).

The quantitative and qualitative instruments and their phase of implementation are summarised in Table 8.

Table 8 – Data Collection Instruments and Implementation Phase				
Pre-Initiative	During Initiative	Post-Initiative		
Pre-Questionnaires Classroom Observations	Classroom Observations	Post-Questionnaires Focus Groups Teacher Interviews		

Any tool that seeks to measure something must be able to produce results that are both accurate and valid (Reid, 2006), this is particularly difficult when it comes to measuring a latent variable. Kind *et al.* (2007, p. 874-875) and Osborne *et al.* (2003, p. 1055-1059) outline five main methods of measuring attitude, opinion or perceptions and their various advantages and disadvantages:

- *Preference ranking:* This is a straightforward process that merely asks participants to rank their school subjects or topics of an individual subject in order of favouritism.
- *Attitude scales:* There are several attitude scales, differential, semantic differential and summated scales.

- *Interest inventories:* Participants are provided with a list of items and they must select the ones they find interesting.
- Subject enrolment: Gathering data on the number of enrolments for individual subjects
- Qualitative methods: Gathering attitudes and opinions through interviews and focus groups.

Attitude scales are seen to be the most widely implemented method of measurement, particularly in educational research (Kind *et al.*, 2007; Reid, 2006; Osborne *et al.*, 2003). These have the ability to ask multiple questions that address the same construct, therefore increasing the reliability of results (Kind *et al.*, 2007). This method was identified as suitable for this study and is therefore discussed in more detail later on in the chapter.

Qualitative methods have been discussed in preceding sections and play a vital role in this study. Where quantitative attitude scales are used in this study to probe students' overall perceptions, qualitative methods are used to investigate the experiences behind such perceptions and how their experiences of day-to-day physics lessons compare with the OAs. They are also used to triangulate data. Given that this research follows a constructivist epistemology, it is highly interpretive. If multiple methods can demonstrate similar evidence, the reliability and validity of the data is stronger. This is addressed more sufficiently in Chapter 8.

When assembling data collection instruments the researcher began by reviewing existing, prevalidated instruments as no sense was seen in reinventing the wheel should appropriate tools already exist. However, given the niche focus of this study, no existing instruments were identified as effective in addressing the specific objectives and collecting the data that was sought in its entirety.

The following sections therefore provide detail regarding the process of constructing this study's data collection instruments. The specific design for the questionnaire, in terms of its design and piloting, the design of the observation protocol and the design of the focus groups and interviews are described. Issues with validity and reliability is touched upon but is discussed in full in Chapter 8.

7.2 Perceptions of Physics Questionnaire

For most educational research, the main objective of a questionnaire is to discover something specific about the sample population at a given time (Verma and Mallick, 1999).

The Perceptions of Physics Questionnaire (PPQ) was the guiding tool in investigating students' experiences of their physics lessons and of the educational OAs. The PPQ was implemented pre- and post-OA implementation to allow for comparison between the two contexts. The pre-PPQ summarised students' general perceptions towards physics in school. The post-PPQ then applied the same method to assess students' perceptions and learning experiences of the OAs.

The foundations of one's perception encompasses a wide range of constructs and characteristics that are influenced by an infinite number of environmental cues. Therefore no single questionnaire can accurately measure every single aspect. It also would not necessarily be advantageous if one could. It is considered that keeping the number of items in a questionnaire to a minimum is beneficial both in terms of time constraints and quality of responses. Therefore, researchers tend to focus on a number of aspects relating to latent construct that they are interested in.

In the case of this study, the pre- and post-questionnaires were designed to focus on six key constructs, largely based on the review of literature in Chapter 2 and the nature of the OAs developed in this study:

- General interest towards physics lessons
- Confidence in physics lessons
- Aspirations in physics
- Interest in astronomy
- The pedagogical experiences of students
- Perceived relevance of their physics curriculum

These were deemed important attributes to students' perceptions towards physics and whether or not astronomy offers a suitable context.

However, due to the focus on perception rather than attitude, and a comparison rather than change, the aspiration in physics construct was only included in the pre-PPQ in order to gain an insight into students predispositions. Such aspirations were not anticipated to change over just one lesson and are not a focus of the OAs and were consequently not relevant in the post-PPQ.

The full pre- and post-PPQ can be found in Appendices 8 and 9.

7.2.1 Types of Questions

During the design process of the PPQ, consideration was given to how the response data was to be analysed and applied once collected.

There are two broad categories of questions; closed and open. When answering a closed question, participants must select an answer from a set of defined responses. Open questions however grant respondents the opportunity to write a free account of their opinions in their own terms (Cohen *et al.*, 2011). Both of these question formats are used in the PPQ. However, there are many sub-categories under these umbrella terms, all with differing advantages and limitations.

Although primarily composed of Likert-scales, the PPQ also included dichotomous, semantic differential and open questions. A 42-item Likert scale (35 in post-PPQ), made up part one of the PPQ and part two was a combination of these other question formats. Before breaking down the PPQ into

each of its component questions, each of the question formats that are used are now outlined and reviewed in terms of their advantages and limitations.

Dichotomous questions are used when you want participants to choose one of two options. The most common form of this question is the simple yes or no question. The main advantage of dichotomous questions is that respondents are obligated to make a clear cut decision, forcing their hand one way or another. However, not surprisingly, they do not provide particularly rich data. Only one dichotomous question is used in the post-PPQ and the question that follows asks respondent to expand on their answer (see question 5 in Appendix 9).

Semantic differential and Likert items are both forms of rating scales (Cohen *et al.*, 2011). Semantic differential scales involve asking participants a question which they must respond to on a scale with one adjective at one end of the scale and its direct antonym at the opposite end. Respondents then decide where their opinion lies between the two extremes. For example, 'I find my physics lessons...', with Easy at one end of the scale and Difficult on the other end.

Summated Likert scales were developed in 1932 by Rensis Likert as a technique to measure peoples' attitudes. Since then they have developed dramatically and Likert scales are now also used to measure aspects such as emotional states, personal desires, personalities and experiences (Spector, 1992). All items within a Likert scale are provided as statements, the participant must then choose one of several response choices that best reflects their disposition to the item; i.e. a level of agreement or frequency of occurrence. There are typically between four and seven response choices, but it is essential that altogether, the choices exhaust the range of possible answers the participant may want to give (Cohen *et al.*, 2011). Some researchers will also include "*I don't know*" or "*no opinion*" options.

Likert scales are widespread throughout educational literature. One of their strongest advantages is that they offer the ability for the researcher to combine a quantitative measurement with qualitative data (Cohen *et al.*, 2011). Many forms of data which may not appear to be quantitative can be collected with Likert-scales (and other instruments) that allow responses to be later quantified. This is a common technique in educational research and is often referred to as a conversion design (Johnson and Onwuegbuzie, 2004). If developed and validated properly, Likert-scales are able to measure various psychological traits of individuals just through analysis of their responses to particular items and therefore can also hold strong psychometric properties (Spector, 1992).

However it is acknowledged that they are not without their limitations. Qualitative response options are notoriously less reliable than quantitative options as answers are based upon participant interpretation. For example, they are subjective, so one participant's perception of 'strongly agree' might be another participant's perception of 'agree'. However, summated scales are more reliable than yes-no questions which are restricted to just two levels of measurement (Spector, 1992). It is fair

to assume that people who respond with a 4 on a 5-point Likert scale will be more comparable in their opinions than those who respond 'yes' on a yes-no question.

Open questions are incorporated in order to expand on the participants' responses to the closed questions. Although termed 'open', some of the questions could still be considered closed to a certain extent as they are limited in the responses that can be provided. One example is, "what is your favourite subject as school?". Students can only respond within the limits of the subjects they are taught. Open questions are generally still relatively specific in the question they ask, therefore the respondent can only really express their opinion in that particular area. Having a final question that asks "do you have any other comments?" can ensure that participants have the opportunity to report on any other thoughts they have not been able to share in the rest of the questionnaire.

7.2.2 Designing the PPQ

Unsurprisingly, due to the uniqueness of this study no single questionnaire was found that addressed all six of the targeted constructs, and so an existing questionnaire could not be employed in full. However, instead of generating an entirely new questionnaire, items were adopted from several existing questionnaires that had been previously validated and demonstrated to work with this audience.

Relevant items were adopted from four educational research studies aimed to measure various attributes of students' attitude to science. These were: Kind *et al.* (2007), Bartlett *et al.* (2018), Aldridge *et al.* (2012) and the National Centre for Education Statistics (NCES, 2012). Adopted items were then combined with some original items, to create a set of questions that collectively, covered all aspects of the research and the objectives (Verma and Mallick, 1999). Of the 42 items overall, 17 were original and 25 were adopted from existing studies, 12 from Kind *et al.* (2007), six from Bartlett *et al.* (2018), six from Aldridge *et al.* (2012) and one from NCES (2012).

The questionnaire of Kind *et al.* (2007) sought to address six constructs: practical work in science, science outside of school, importance of science, self-concept in science, future participation in science and attitude to school in general. Kind *et al.* also adopted items from existing questionnaires including the Relevance of Science Education (ROSE) questionnaire, the 2003 PISA questionnaire and the attitude to science questionnaire developed by Pell and Jarvis (2001).

The items adopted from Bartlett *et al.* (2018) came from their newly validated interest in astronomy scale aimed to be embedded into general attitudes to science scales in order to assess the impact of astronomy-based initiative projects in education. This instrument was validated through responses from 429 students aged 14-17 in Australia.

Aldridge *et al.* (2012) developed and validated a new instrument they designed with a purpose to provide teachers with feedback regarding students' perceptions of the classroom climate. They validated their questionnaire using data from 2,043 16 to 18 year old students in Australia.

The US National Assessment of Educational Progress (NAEP) implements questionnaires to students via the National Centre for Education Statistics (NCES), teachers and school administrators on a voluntary basis. The data collected is used by practitioners, policymakers and researchers in order improve education throughout the United States (NAEP, 2018). The wordings of some items adopted from here were adjusted slightly in order to coincide with the objectives of this study.

Part 1 of the PPQ was therefore a 42-item Likert scale aimed to measure students' perceptions around their everyday physics lessons and towards the implemented OAs. Students were asked to respond to each statement using a five-point scale with the options of 1 =Strongly Disagree, 2 =Disagree, 3 = Neither and 4 =Agree and 5 =Strongly Agree for 36 of the items and 1 =Never, 2 =Rarely, 3 =Sometimes, 4 =Often and 5 =Always, for six items.

Kind *et al.* (2007, p. 6) propose four guidelines that should be followed in order to construct an effective attitude measure. These points are detailed below and discussed in terms of how they are each addressed in this study.

1. Clear descriptions of the constructs intended to be measured

The scale focussed on six key constructs that, based on the literature reviewed in Chapter 2 were deemed important and suitable attributes to students' perceptions of their learning experiences, and additional sixth construct based around the context of the implemented OAs. The scales were:

- General Interest
- Confidence
- Future Aspirations in Physics
- Interest in Astronomy
- Pedagogical Experiences
- Perceived Relevance

It is acknowledged that these constructs are not exhaustive and additional factors that are influential to students' learning experiences in physics exist. There are existing questionnaires that have been developed to measure students' attitudes, some of which address these additional constructs but will not address others. However, these six constructs were chosen to be of highest value and relevance to this study and so formed the focus of the PPQ.

General Interest is a frequent construct in the research around attitudes to science (Osborne *et al.*, 2003). It is well understood that if someone holds an interest towards something, they will generally report positive experiences and hold a positive attitude towards it.

Extensive literature indicates that students with high confidence in a subject, report greater interest and enjoyment in that subject than students with low confidence. Literature is also indicative of students perceiving physics as difficult and is therefore considered to be a valuable measure.

Although this study did not anticipate that such short-term interventions will completely alter students' career path, the researcher deemed it important to know students' aspirations at the baseline. Particularly as random sampling was not used, this helped to identify how the perceptions of student participants in this study compared with those reported in the wider literature.

The construct Interest in Astronomy focused solely at the context of this study. It was used to investigate whether or not the context of astronomy was an effective point of engagement for triggering student interest and positive learning experiences. This construct was heavily based upon the work of Bartlett *et al.* (2018) who developed and validated an 'Interest in Astronomy' scale. Given the researcher's role, had this validation phase been carried out sooner it is more likely the researcher would have adopted the entire construct. However, as this study was already underway and validation data for the other scale had only been collected from Australian students, a separate scale was produced.

Chapter 3 discussed the support for inquiry-based learning in science education and its positive effects on students' learning experiences. The construct regarding pedagogical experiences was therefore aimed to measure students' perceptions in this study. The items in this scale investigated whether students felt they experienced IBL both in their day-to-day learning (pre-PPQ) and in the OA (post-PPQ). But most importantly, it investigated whether it encouraged positive learning experiences.

The terminology used throughout the adopted items to describe these learning experiences was 'practical work'. This was replaced with 'investigation work' in the PPQ in order to reduce confusion around the definition of practical work. As the activities involved either using an online simulator or remote, robotic telescopes, it was anticipated that there was a risk that this would not be regarded strictly as 'practical' as it didn't require any physical equipment in the classroom other than computers.

Feedback from a focus group carried out in the questionnaire validation phase (see Chapter 7.2.3) also indicated that practical work is perceived as a more 'hands-on' activity which would not necessarily be considered to include the use of robotic telescope data or computer simulations. Furthermore, it is considered by the researcher that 'investigation' is more reflective of inquiry-based learning than 'practical'.

Problems with student perceptions towards the relevance and usefulness of their physics curriculum was also highlighted in Chapter 2 and is considered as another contributing factor to students' learning experiences. In consideration of this and given that the Department for Education in England prescribed that science education should be delivered to students with a clear relevance to their everyday lives, it was included as a construct in this study.

This construct was slightly different to the other five in that, although responses remained on a 5-point scale, here students responded with a level of frequency rather than agreement. Each item in the scale began with the phrase "in my physics lessons I am able to…".

2. Care must be taken and justification must be given when combining separate constructs within one scale

Care also had to be taken in deciding on how many items to include in the questionnaire. Generally, the more items included to measure a construct, the higher the reliability (Spector, 1992). This is because by using a higher number of items that each addresses a construct; you are more likely to have covered every influential and underpinning aspect relating to that given construct. This is particularly important when seeking to measure a latent construct, such as perception and experience. However, it is also emphasised by Kind *et al.* (2007) that a balance between an array of meaningful and relevant statements and limiting the number of items on a questionnaire is crucial. It is generally understood that people do not appreciate long questionnaires, therefore if they get fed up, they are more likely to provide more negative responses, regardless of their relevant perception. Bartlett *et al.* (2018) describe this as a parsimonious solution. The more representative and encompassing the questions are of the entire construct, the less error is seen in measurements. Maintaining that measures are relatively short however, is effective in reducing boredom or fatigue in participants that will typically lead to response bias (Hinkin *et al.*, 1997).

Although various recommendations are made, it is apparent that there is no single, universally accepted rule regarding the optimum number of items. Hinkin *et al.* (1997) argue that adequate quality can be achieved with just four to six items for one construct. Clark and Watson (1995) however, prescribe that the items should be broader than the researcher's view of the target construct and if anything, should overcompensate. Their argument is that psychometric analyses can identify non-essential or irrelevant items to be later removed, but cannot commend items to add to the scale that were not already included. This latter approach was also employed by Bartlett *et al.* (2018).

In response, it was decided that each construct should have between six and eight items. This was to ensure that each construct was sufficiently addressed and resulted in no less than 30 items in total, but also did not exceed 50 items, therefore not demanding too much time from participants and risking response bias.

- 3. Reliability must be demonstrated through confirmation of internal consistency and unidimensionality
- 4. Validity must be demonstrated through multiple methods including psychometric techniques.

To demonstrate points 3 and 4, psychometric analyses are carried out in the data analysis phase, in order to confirm the unidimensionality and internal consistency of items and constructs. Here, the former describes the existence of a single latent construct that underlies a set of measures (Gardner, 1975), in this case students' attitudes towards physics. The latter implies that this set of measures (or items), correlate with one another and measure the same aspect.

As in this instance there are six constructs that are used to address students' perceptions, it is essential the items that make up each of these six constructs is unidimensional, but given that such perceptions are being inferred by multiple constructs, by definition the questionnaire in full, is multidimensional. Unidimensionality and internal consistency of items within each construct were not assessed until all data collection was complete.

Table 9 – Number of Items Contributing to each Construct in the PPQ				
Construct	Number of Items			
General Interest	7			
Confidence	7			
Future Aspirations in Physics (pre-PPQ only)	7			
Interest in Astronomy and Space Science	8			
Pedagogical Experiences	7			
Perceived Relevance	6			
Total	42			

A summary of the constructs and the number of the component items is provided in Table 9.

7.2.3 Trialling and Refining the PPQ

The second part of the PPQ was designed in order to gather information from students that could not be retrieved from the Likert-scale items in part 1. The construction of this second part of the PPQ stemmed from the feedback from the pilot test of the Likert-scale (part 1) of the PPQ where it became clear that certain areas needed greater focus and there were also occasions where participants in the trial reflected on important perceptions of their physics lessons that were not explored by the PPQ.

The PPQ was piloted with a group of individuals from within the target population in the form of a focus group. Participants included six year 10 students, three male and three female. As the post-PPQ asked participants to reflect on experiences of the OAs, these students just completed the pre-PPQ that

focuses on their experiences of school physics. The primary purpose of this pilot was to identify the suitability of the questionnaire to the target population, any missing features and any problems or misinterpretations. Hunt *et al.* (1982, p. 270) plainly state that the ultimate aim of piloting is to recognise:

"Will the instrument provide data of sufficient quality and quantity to satisfy the objectives of the research project?"

Piloting is the only way to assess whether respondents have any problems completing the questionnaire, prior to official implementation (Presser *et al.*, 2004).

The remainder of this section details outcomes of the PPQ trial and items that were added to part 2 of the PPQ as a result. The purpose of each question in part 2 is described and justification from pilot feedback is provided.

Participants in the trial were given an initial briefing where they were explained what was expected of them. As recommended by Gall *et al.* (1996), participants were instructed not only to complete the questionnaire, but also to annotate it along the way if they had any criticisms about the questionnaire or any recommendations for improvement. Completion of the questionnaire was then followed by a focus group to allow the researcher to probe students' responses and identify how each question or item was interpreted and what students reflected on when forming their response. Students were strongly encouraged to express any criticisms or recommendations that they held with regard to the structure of the PPQ.

Much insight into the appropriateness and effectiveness of the questionnaire was revealed from students' responses alone. Nonetheless, after completing the questionnaire, students were first asked, what do you think is the general purpose of this questionnaire? This prompted the following responses:

"whether we like physics lessons" "what helps us to learn in physics" "how we learn physics" "what parts of physics we like and don't like"

"whether the syllabus is useful, how you best learn from the syllabus, best method of learning"

These responses were considered to be very promising and reflected that the questionnaire successfully portrayed its intention and what students needed to consider in their answers, thus demonstrating construct validity.

To follow this, participants were then asked, 'what were you thinking of when answering the questions?' Students reflected that they had been thinking about the classroom environment, the

teacher, difficulties and challenges they face, individual topics in physics and also the subject of physics as a whole. Again, the researcher was happy with these responses and felt that the questionnaire had successfully demonstrated its purpose to the target audience, and prompted them to reflect on relevant experiences.

Participants were asked about aspects of the questionnaire as a whole, such as the length, format, and sequencing of questions (Hunt *et al.*, 1982). This yielded positive responses and students indicated that the flow of the questionnaire seemed appropriate.

Finally, students were asked if there were any experiences of physics lessons that they felt they weren't able to express in the questionnaire.

One student mentioned that they were not able to discuss the teacher's influence on their experiences in physics. Another student expressed, and other students nodded in agreement, that there wasn't much discussion of the overall classroom environment. Both of these points were considered to be valuable and important elements within a classroom.

On reflection of these comments, the researcher felt that it was important to use the questionnaire to gather information from students that could not be gathered from classroom observations or focus groups. In order to obtain a solution that maintains both a parsimonious questionnaire that is also rich in data, certain elements had to be prioritised above others. Classroom environment was considered by the researcher to be more readily measured from student and teacher behaviours within the classroom and communications that take place, thus available through classroom observation records. Where questions regarding the classroom environment would provide a broad understanding of students' perceptions, classroom observations can directly witness students' engagement in the lesson, engagement with their peers and the teacher and also at what specific point such behaviours were seen, i.e. what part of the activity and does not summarise the classroom environment as a whole as would be reflected from questionnaire responses.

One student expressed that for some items, a scale of agreement was not well suited to the items and that an answer of "sometimes" would have been more appropriate. The researcher asked if a scale of frequency would be more suitable in these cases, to which this student confirmed and other students nodded in agreement. It was these items that were subsequently allocated a frequency scale ranging from 1 =Never to 5 =Always, instead of an agreement scale (all items were initially presented on a level of agreement scale).

A number of additional items were also added to the existing Likert-scale constructs in order to more fully address relevant aspects of students' perceptions that were reported in feedback during the pilot.

A question that asked students if they could recommend any improvements to the activity was also added to part 2 of the post-PPQ. Though there was already a question that asks students what they

liked least, it was felt that this would force students to limit their response to just one aspect. However, when asked about what improvements they would recommend, though they might reflect on what they claimed to like least, they could also add additional ideas if they so wished. This question could not only reflect on the OAs themselves and guide future developments, but could also reflect implementation. These comments could therefore be provided to teachers who use the OAs in the future when deciding how they want to implement the OAs in their classroom.

On reflection of the pilot study, the researcher considered that it would be impossible to ensure that the questionnaire successfully covered every aspect of physics education and the initiative activity that students may wish to voice. As a compromise a final question was incorporated that asked participants to detail any other comments they have about their physics lessons (pre) or the OA (post). This gave participants the opportunity to portray any opinions they held but were not addressed in any of the previous questions. This final item is used to combat the issue of 'not knowing what you're not asking' that is accompanied with questionnaire data.

This trial phase of the PPQ demonstrated to be very useful in finalising the questionnaire and gaining a clearer understanding of the target audience. The final version of the PPQ was considered to address all areas of interest to the researcher and also gave students the opportunity to express their opinions fully.

The final version of the PPQ that was implemented to students in this study can be found in Appendix 8 (pre) and 9 (post). The process of the validation following data collection was parallel to that of Bartlett *et al.* (2018) and so is not reported here.

7.3 Classroom Observations

Observations are a form of naturally occurring qualitative data (Ritchie, 2003) and can take a variety of formats such as facts, events and behaviours (Cohen *et al.*, 2011). Observations provide genuine, *live* and direct information of social behaviour in a natural environment rather than *recounted* information that is provided in follow-up questionnaires and interviews (Cohen *et al.*, 2011; Ritchie, 2003). As case studies typically take place in a natural environment they provide an ideal opportunity for direct, unprompted observations of participants (Yin, 2009). Support for this is provided by Robson (1993, p. 205) who highlights that in smaller scale projects, participant observation is particularly useful when a key motivation is to *"find out what is going on"*. This ties back to the nature of this study and the point by Cohen *et al.* (2011) highlighted earlier where interest lies in how things work, for whom and under what given circumstances. Participant observation also permits the researcher to immerse herself in the research environment and experience scenarios with participants. It is considered that observational data is highly valuable in this case study design in terms of

exploring how students engage and interact with both their day-to-day physics lessons and the implemented OAs and react to the context and process as a whole. Guba and Lincoln (1981) express that laboratory-style experiments are not appropriate for the study of social groups as they are insufficient in accounting for complex human behaviour. In scenarios where human behaviour is governed by attitudes and beliefs, the most valuable instrument to a researcher is a careful observer, who has the ability to analyse and assess the scenarios from direct experience.

Due to the naturally occurring transfer of data, observations not only reinforce other data collection methods which can be triangulated in order to verify the authenticity (Cohen *et al.*, 2011; Robson, 1993), but they also provide an insight into areas that may have been unconsciously missed or overlooked in the quantitative phase (Cohen *et al.*, 2011) or simply were not possible to measure.

Classroom observations provide a direct awareness and perspective of any issues with either the OAs themselves or their implementation. Equally, aspects that work particularly well and prompt high levels of engagement, or those that do not and prompt low levels of engagement from the students can be naturally recognised (Yin, 2009). Observation measures can also help to identify and account for courtesy bias in the questionnaire results. This is seen when inconsistencies are displayed between questionnaire responses and observation data due to participants' behaving differently to how they claim to (Robson, 1993).

However, observation techniques are not without their limitations. Osborne *et al.* (2003) point out that often behaviour is influenced by *other* attitudes that may be held more greatly by an individual. For example, a student may not have a particularly negative attitude towards physics but may display one in their behaviour in order to conform to their peers or to fit into a stereotype. To the individual, an apparent mutual attitude and similarity with peers may be more imperative than portraying their true attitude towards a subject. Influence of peers is particularly evident among high school students (Chen, 1997; Temitope and Christy, 2015). This is considered to be most problematic among younger participants in this study, year 8 and 9 students. At the point of year 10 and above, the researcher is of the belief that students have already established friendships and are considered to be under less pressure to conform to social stereotypes that are not analogous to their own internal beliefs.

The unintentional influence of the observer can also be problematic. Just the presence of an outside observer could often lead to both the teacher and students changing their normal behaviour and consequently changing the observed environment (Gall *et al.*, 1996). As this aspect is directly relevant to matters of validity, it is addressed fully in Chapter 8.

According to Gold (1958) and Cohen *et al.* (2011), the role of the researcher in this study was an observer-as-participant. Although not a complete member of the group, she contributed low levels of participation in the group's activities. More specifically, the researcher was clearly not a student in the class, nor the teacher, but occasionally interacted and communicate with students during the study.

The observational variables are both descriptive and inferential (Gall *et al.*, 1996) where the former requires minimal inference from the observer and the latter requires the researcher to translate observations of actions and behaviours into interpretations and conclusions. This approach is more formal than the participant-as-observer, where the researcher is much more involved with the group and risk becoming too immersed in the group activity causing the researcher to stray from the observation objectives and could also influence the behaviour of participants. Robson (1993) argues that the correct variables are observed in slightly more formal observations, consequently promoting higher validity and reliability than the more informal, unstructured methods.

However, Gold (1958) highlights that a mere, brief contact with participants stands as a limitation of this role, as it becomes more likely that this unfamiliarity will lead to misinterpretations and misunderstanding. This is another advantage of following a mixed method approach, data collected from the PPQ, classroom observations and focus groups can all be triangulated to yield more sound interpretations. It is also complimentary to a constructivist approach as it can help to understand why certain events occur and under what circumstances (Verma and Mallick, 1999). The researcher is thus able to apply her own experiences within the environment in order to 'read' and interpret the information that is provided by the participants (Verma and Mallick, 1999, p. 29). This permits an ability to explore the larger picture of the given environment and its context.

Despite the drawbacks, Gall *et al.* (1996) maintain that the advantages of classroom observations, far exceed the limitations and provide a valuable evidence base. So long as a descriptive narrative, explanation and justification is provided for all inferences made by the researcher, the reader can follow and understand such interpretations. The researcher's application of reflexivity also aids this process.

7.3.1 Developing the Observation Protocol

Though the research is constructivist in nature and classroom observations are intended to be more open in terms of data collection, a semi-structured protocol was considered to be helpful in maintaining a focus (Cohen *et al.*, 2011). This method is employed where the researcher has a general agenda of behaviours, interactions and communications to look for and there is a general structure and direction for observation (Robson, 1993).

Existing protocols and instruments for classroom observations were initially searched. This revealed many standard observation protocols that had been previously developed and validated, but as was seen with existing questionnaires, they all varied in which aspects of a classroom they apply to and also whom they apply to. Each tool will have been developed with a particular focus and research objective which varies to some extent from those of this study.

What became immediately apparent was the proportion of protocols that appeared to focus primarily on the teacher and their instruction rather than the students. These are generally used for performance management purposes of teachers and their abundance is likely due to accountability pressures that leads schools to be more concerned with performance than experience. Those that had a larger focus on students were very much quantitative-based, typically following a very prescriptive, checklist format. That is, a single student or group of students are observed and the observer has a list of events which they either tick to confirm that such events happened, or they indicate the frequency of each event.

Palpably missing from these existing tools were specific measures to explain the evidence for such engagement and how it was identified. The researcher considers that a simple checklist offers little more depth than a questionnaire and is more typical of a positivistic stance.

The protocols typically addressed are described by Gall *et al.* (1996, p. 332) as descriptive observational variables, with little attention to inferential or evaluative variables. Although the latter two are often developed sequential to the descriptive variables, many protocols lacked the detail to reflect sufficiently on why such aspects and behaviours were recorded.

This approach was considered to be far too restrictive for the nature of this study and the specific purpose of classroom observations. The objective of this research was to identify specifically *how* particular learning experiences are triggered, not in a broad sense of the activity as a whole but by breaking down the activity into separate parts. This also reflects a reason as to why duration and frequency-count techniques were considered inappropriate for the research as the intention was not to count occurrences of positive experiences among students but to identify what led to them and how they were fostered.

In response, all existing protocols that were found were rejected for use in this study. This meant a new protocol had to be developed, the construction of which is now discussed.

Wragg (2012, p. 105) sets out "*twenty questions about the classroom*" that encompass a variety of areas and therefore not all of them are applicable to this study. Those that do are as follows:

- What do pupils learn; what tasks do they engage in, and with what degree of involvement and success?
- Is there coherence in the curriculum [OA content], or does it appear disjointed and incoherent from the child's point of view?
- What happens when children work in small groups; what kind of assignments are undertaken; who decides what; are the groups collaborative?

Gall *et al.*, (1996, p. 333) classify observation recordings into four key techniques; duration, the duration over which a specific behaviour takes place; frequency-count, the number of times a

particular behaviour occurs; interval, observing an individual's behaviour at given intervals; and continuous, recording all behaviours of one or a group of individuals over a specified observation interval. Where frequency-count and to an extent, continuous were most commonly seen among the literature, a combination of continuous and interval techniques were employed in this study.

The researcher developed two tools to use in classroom observations, a summary of demographics and delivery and the protocol itself.

Demographics covered aspects such as the time and date of measurement, number of students in attendance and gender breakdown, gender of teacher and seating plan. Delivery provided a summary of how the initiative activities were implemented. It included details regarding pre-planning and preparation, what students were provided with and a broad description of delivery. The demographics and delivery aspect was used to provide a context to each of the case environments, it also set the scene for the more detailed observations provided by the protocol.

The protocol itself was made up of ten main areas for observation as well as a running observation record and an area for the researcher's interpretations. These were not as constrained as those found in the literature but encompassed broader areas that could extend down a number of directions without straying too much from the focus of the observations. This lack of constraint was based on the inductive approach of the study and the intention for goal-free data collection.

The ten points of observation are provided below. Explanations of each can be found in Appendix 10.

- 1/. Unplanned Activities
- 2/. Individual, Paired or Group-Work
- 3/. Capability of Set Tasks
- 4/. Communications
- 5/. Time Optimisation
- 6/. Body Language
- 7/. Behaviour of Teacher
- 8/. Behaviour of Students
- 9/. Comments from Teacher
- 10/. Comments from Students

DeWalt and DeWalt (2002) also recommend maintaining a running observation record, in order to ensure as many aspects of an environment are captured, even if not appearing to apply to any of the other protocol areas. This final section was therefore intended to record all observed events and permitted the application of goal-free data collection (Scriven, 1991). It is considered that while in the moment, where such events should be placed in terms of the protocol main areas may not be entirely

clear. Some observations may also seem irrelevant at the time but should still be noted. Observations listed in this area were therefore to be reflected on in post-observation in order to better understand them and where they were relevant.

Given that school classes can be large, with up to 30 students, classrooms were divided into several of groups, typically based around the seating plan. Each group was observed over intervals for a particular period of time. Though, it was decided that anytime where something external to the current group of focus drew the attention of the researcher, this should not be ignored and attention could shift in response if it was felt necessary.

Wolcott (2002) suggests when one is unsure on what to focus on, the researcher should turn to what they happen to be attending to at that time, specifically what drew their attention and why. They should also continuously review what it is they are looking for, what is necessary and whether it is presented in the particular scenario they are observing at that point in time.

This final section of the protocol, researcher's interpretations, had a slightly different purpose to the other 11. This section was not used to record direct observations, but was used to reflect on observed events or behaviours and record how they were interpreted by the researcher and what they were considered to imply. For example, if a student was seen to ask additional questions about the activity they were completing, this could be interpreted to reveal interest and curiosity in the topic. Similarly, should a student be seen to be leaning back in their chair, gazing out of the window, this would likely be interpreted as disengagement and disinterest in the task.

Unfortunately, it was not possible to trial the protocol before the data collection began and so the first observation in the study also had to act as a test for the protocol which did reveal some difficulties in implementation of the protocol.

What was immediately evident was the difficulty in section 12, the Researcher's Interpretations. The process of recording interpretations meant less time was available for both recording and observing live observations. This meant that important events or behaviours were at risk of not being seen. Given that the first implementation involved a class of 26 students, observation time had to be split across groups of students, reducing the average time spent observing each individual student.

On reflection of this it was decided that this section would be better completed, subsequent to completing the classroom observation. In future implementations, this section was only completed after observations had been completed. The focus of actual observation was to record only what was seen, not what was interpreted. Instead, interpretations would be made immediately after returning from the venue of implementation.

This reflection immediately after implementation was considered to be one of the most valuable points for observational data. At this point, observations were still fresh and recall was most efficient.

As many observations were taken in short hand, all observation notes were expanded to attribute additional context and elaboration. This was done for almost every observation note made, and following this, the researcher then made her interpretations.

However, observation records were not left after the day of implementation and were returned to both a week and a month later. This proved to be extremely valuable as additional occurrences were recalled and various interpretations were considered, based on time away from the data. This also incorporated processes of reflexivity whereby the researcher's experiences increased over time in accordance with additional implementation events (Cohen *et al.*, 2011). This therefore influenced the researcher's interpretations and developed her understanding of classroom environments.

This process of reviewing the classroom observation data is illustrated in Figure 6.



Figure 6 - Flow Chart of the Review Process of Classroom Observations for each Group.

Templates for the classroom observations demographics and delivery and protocol can be found in Appendix 10.

7.4 Student Focus Groups

Turning now to the qualitative sequential method of data collection, semi-structured focus groups were only carried out subsequent to implementation lessons and were informed by PPQ and observation data. Semi-structured focus groups are commonly seen in mixed method research and are used to accompany other data collection instruments in order to triangulate data (Morgan and Spanish 1984; Vaughn *et al.*, 1999) and also expand on data from quantitative phases (Creswell, 2009).

Focus groups involve directing a series of questions to a group of participants who have been gathered specifically (Gall *et al.*, 1996; Punch 2009) to discuss a particular topic with each other and with the researcher (Morgan and Spanish, 1984). Focus groups are recognised as a method of data collection that allows participants to express their internal views more freely and with less limitations and restrictions of a questionnaire. They provide additional, more in-depth information that cannot be obtained through single-participant interviews and questionnaires alone (Bailey and Slater, 2005; Finch and Lewis, 2003). This is because interaction is not only with the researcher but also with each other, participants are able to reflect on one another's responses and develop and build on their own views, thus widening the group discussion (Gall *et al.*, 1996; Kitzinger, 1995; Finch and Lewis, 2003). This is described by Hess (1968, p. 195) as *snowballing* and is not possible to achieve through questionnaire data as each participant has their own individual questionnaire which they typically complete by themselves with no interaction with other participants.

The focus groups were considered to be semi-structured in the sense that the researcher asked a series of key questions in order to gather particular information that relates to the research questions (Arthur and Nazroo, 2003), but conversation was not restricted to such questions and was able to naturally progress in other directions, within reason.

Similar to the observation protocol, a semi-structured approach to focus groups was chosen over a more open-ended format to provide security in ensuring that the key aspects set out to be explored through the focus groups were followed and there was less risk of conversations digressing towards topics less relevant to the study. Although focus groups are considered to be a more open method of data collection that have the ability to take a number of different directions from participants' responses, they should still be structured in a way that ensures all participants across groups have similar experiences (Arthur and Nazroo, 2003). That is, where multiple focus groups are carried out, the researcher should gather similar information across all of them rather than ask entirely different questions. Although students may reflect different experiences, these should be with regard to parallel circumstances.

Focus group data was supplementary to the PPQ and classroom observations and were implemented in order to allow participants to expand on the views that were brought to light in their questionnaire responses, thus providing assurance that the right questions are being asked and the right interpretations were being made. The key questions that guided the focus groups were therefore not designed until preliminary results were gathered from the PPQ and classroom observations. This meant that anything interesting that came to light in these former data collection procedures could be followed up and interrogated further through the focus groups. This process provided the sequential explanatory nature of the study.

Although predefined questions were prepared ahead of conducting the focus groups, they followed a relatively conversational format where sub-questions were often guided by responses to these initial key questions. This ensured that participant responses were more genuine and emerged more naturally, allowing the researcher to further probe the perceptions of the participants and allow them to elaborate on their own views (Gall *et al.*, 1996; Vaughn *et al.*, 1996).

Some argue that in group interviews, participants may tend to express the same views regardless of their true opinions in wanting to reach a group consensus (Cannell *et al.*, 1989). In an attempt to overcome this, during the initial brief, the research clearly stated that "there are no right or wrong answers" and that "different and conflicting ideas between you are welcomed". Students were also reminded that these interviews were anonymous and any comments they made would not be echoed back to their teacher. This was to encourage honest opinions and deter students from answering in a biased manner.

Though the focus groups were designed to follow up on evidence from PPQ and observational data and triangulate results in order to increase validity, it was also vital to ensure that a greater level of depth was provided in the focus group data and that it did not simply replicate the data yielded from the questionnaires and observations (Verma and Mallick, 1999). The intention was to be able to understand and explain the findings from other data sources (Vaughn *et al.*, 1996). Greene *et al.* (1989, p. 266) describe this method of using data from one source to elaborate and develop data from other sources as '*complementarity*'. Complementarity differs from triangulation in that one source of data *influences* another, rather than remaining entirely independent of one another.

Here the researcher sought to achieve both triangulation and complementarity. Triangulation in the sense that evidence from one source of data collection should reflect similar interpretations from the other two sources but complementarity in the sense that these three sources should not just be limited to confirmations of one another but should guide one another to explore results from one source in greater depth. This allowed the researcher to gain greater understanding of students' experiences and what guides their perceptions.

Ideally, focus groups would have been carried out with each participating group in the study. However, focus groups, unsurprisingly, proved to be the most difficult to organise. Focus groups required students to either be removed from lesson time, break periods or to stay on after school, and students and teachers didn't stand to gain much else from the process since they had already completed the implementation activity.

Two focus groups were successfully arranged for group Navy and Brown in the D2E case study, sadly, despite several conversations and prompts, no focus groups were successfully arranged with any groups from the Hubble case study. It is likely because implementation of the Hubble OA always occurred towards the end of the academic year where the relevant content was taught, and at which point, schools became very exam focussed.

Group Navy and Brown were the first and last groups where implementation occurred in D2E. Students from group Navy therefore offered the opportunity to investigate whether students had carried any longitudinal experiences from the OAs as the interview was carried out 17 months subsequent to implementation. This provided an opportunity for the researcher to reflect on students' recall of the activity. Where their PPQ responses immediately after implementation were positive, the researcher could investigate if this was still the case over a year on and after progressing further through their education.

Students from group Brown had a much shorter period between implementation and the focus group with a gap of just two months. This shorter timeframe meant students could more readily reflect on their experiences and expand on the perceptions they portrayed in the PPQ.

When the difficulty of arranging focus groups became apparent, the researcher saw it beneficial to attempt to conduct interviews with participating teachers. This was successful with one teacher in each case study, Mr Brown from the D2E case and Mrs Yellow from the Hubble case. Teacher interviews had not initially been set out as a priority as focus was attributed to students' perceptions of their learning experiences. However, in light of the difficulties in arranging student focus groups, teachers' insights were able to offer additional data that could be triangulated with other sources and methods.

The demographics of the focus groups with Group Navy and Group Brown and questions the students were asked are provided in in Appendix 11. Though it should be recognised that such questions were used as the general structure of the focus groups, in reality, more questions were asked in response to students' answers and interactions.

7.5 Teacher Interviews

As stated, teacher interviews had not initially been part of the data collection plan and were only designed once implementation had begun when the difficulty of organising focus groups became apparent. However, data from teacher interview was not just used to enrich the data collection and

explore elements of implementation practicalities, but was also used for participant triangulation, in that students' responses and teachers' responses could be cross-checked (see Chapter 8).

There was concern whereby research into the reliability of teacher reports on classroom environment and activity are concerning in that often teachers will reflect more positively on such factors compared to both their students (den Brok *et al.*, 2006; Fisher and Fraser, 1983; Hook and Rosenshine, 1979) and outside observers (Hook and Rosenshine, 1979).

However, an interesting finding from Centra (1973) was seen where although teachers reported more positively on different classroom aspects than their students, what was seen to be parallel was the ordered ranking of the same aspects. That is, where teachers reported more positive scores than their students to aspects a, b and c in the classroom, when placing aspects in order of scores, highest to lowest, the data demonstrated a Spearman rank correlation of 0.77, demonstrating a relatively high similarity. Remarkably, Fraser (1982) performed a similar study measuring student and teachers scores across five dimensions. He also demonstrated a rank correlation of 0.77.

Despite the divergences between student and teacher reports of classroom environment and teachers not guaranteeing a reliable source of evaluation (Hook and Rosenshine, 1979), given that similarities are seen across the ranking of activities, and positive and negative experiences more largely, in conjunction with the notion that this study does not rely solely on teacher reports, it is considered that teachers can provide further valuable evidence.

Ultimately, they are also the gatekeepers to the classroom and it is they who decide on what events take place and are implemented within lessons. Their views of the OAs should therefore not be ignored.

Teacher interviews were also semi-structured based on the same arguments as those for the student focus groups. The questions were largely directed around their perceptions of the educational OAs, the success of implementation, how they believed students responded and the advantages and limitations of the OAs. Questions specific to some of their students were also included depending on the teacher who was interviewed.

Interviews were carried out with Mr Brown from the Down2Earth case study and Mrs Yellow from the Hubble case study. A description of when and where these interviews took place as well as a full listing of the questions that were asked to both teachers are provided in Appendix 12.

7.6 Summary

Chapter 7 has argued the case for each of the sources of data collection employed in this study and how they were selected and designed to most effectively address the aim of this study and its research

questions. Table 10 summarises these instruments, their main objectives and the phase of implementation in which they were utilised.

The PPQ acts as the initial fact finding technique (Verma and Mallick, 1999), where the classroom observations and follow-up focus groups and interviews stood as the explanations, reasons and processes behind this surface level data.

The questionnaires, observation protocol, focus group and interview templates can be found in Appendices 8 to 12.

Table 10 – Summary of Data Collection Instruments					
Instrument Objective		Implementation Phase			
(a)	Questionnaires	Assess students' attitudes and perceptions.	Pre and Post		
(b)	Classroom Observations	Recognise aspects perhaps missed by (a) and assess whether claims from (a), (c) and (d) are reflected in actions.	During		
(c)	Student Focus Groups	Deeper investigation of reasons behind students' perceptions reported in (a) and behaviours witnessed in (b).	Post		
(d)	Teacher Interviews	Investigation of teachers' opinions around students' learning experiences and for participant triangulation of data collected in (a) and (c)	Post		

Table 10 – Summary of Data Collection Instruments

Chapter 8

Addressing Validity and Reliability

Justification for the employment of a sequential explanatory, mixed method, multiple-case study design has been described in the preceding chapters along with details of instrumentation and how each is best suited to address the research questions of this study. However it is recognised that no research design is without its limitations and threats to validity.

Merriam (1988) highlight that in order for an initiative project to influence theory or practice in educational fields, such studies must be trusted by stakeholders. Whether or not a research study can be trusted is described by its validity and reliability and is therefore essential in all fields of research.

Validity describes the truth value of research findings (Guba and Lincoln, 1981) and the extent to which the reader can trust the process of the study and that resulting conclusions are sound. Reliability refers to the consistency of such findings and whether these would be observed again if repeated. Reliability is arguably irrelevant if a study is not valid as there would be no need to justify the reliability of empty and meaningless findings.

As with their philosophical assumptions and approaches, social science researchers, educational or otherwise, are not unanimous in their views of demonstrating research validity and reliability. Two broad approaches are seen to be employed, the use of which depends on the nature of the research. Traditional philosophies of validity and reliability are those of positivistic researchers who view reality objectively and assess research in terms of internal, external and construct validity as well as reliability (Yin, 2009). However, constructivists argue that these definitions were generated in a positivistic tradition and are therefore not applicable to case study or qualitative designs (Yazan, 2015). Merriam (1988) makes the point that where *understanding* is the focus of an investigation, the criteria for trusting such a study will be different to one that focuses on testing a hypothesis. Therefore, different paradigms and epistemological stances should naturally hold different conceptualisations of validity and reliability. Given the nature of this study and the integral role of the researcher, this also presents additional threats that would not be considered by positivists.

However, even among constructivist researchers there are various terminologies. For example Gall *et al.* (1996) contend that research should be assessed in terms of plausibility, authenticity, credibility and relevance, whereas Guba and Lincoln (1981) provide definitions of credibility, applicability,

consistency and confirmability. The following sections explore these different definitions and identify their appropriateness to this study.

8.1 Credibility

Credibility is the accepted alternative for the positivistic definition of internal validity which describes the extent to which we can say that the *experimental stimulus* itself made a significant difference within a given scenario (Campbell, 1957, p. 297). The credibility of constructivist research however describes how much confidence can be attributed to the reported findings or a particular study for the given subjects and context. Put simply, do the findings accurately describe reality?

This is a tough feat for constructivists who must contend with multiple realities, all of which are unique to the participants. Stake (1995, p. 108) emphasises that given the multiple perspectives, there is no way of establishing which perspective and interpretation is the 'best'. Where people are the source of study, i.e. students within a classroom, each of these participants holds their own internal reality and all are true. The researcher must therefore judge the *credibility* of their interpretations with the corresponding sources. This credibility reflects the extent to which the explanations are logical in light of the evidence presented, or whether alternative explanations could hold merit (Hitchcock and Hughes, 1995). The issue of credibility was a primary reason for the application of reflexivity and the importance of reviewing data and interpretations as the researcher's experience and understanding developed. This was also in line with Stake's recommendation (1995) of asking whether both a comprehensive and accurate description of the case has been provided in the data and whether interpretations are developed in the way they should be (Yazan, 2015).

Guba and Lincoln (1981) acknowledge that distortions can arise from biases not only from the researcher but also the participants. Chapter 6.1 set out the initial biases and predispositions of the researcher in a first step to address this issue and highlighted how reflexivity is applied throughout the research process. However the researcher could also go into a classroom environment with expectations of the participants or of the OAs or biases could present themselves in the analysis stage. With an awareness of this, the researcher approached the classroom observations with an aim to simply 'write what is seen'. Chapter 7 explained that no interpretations would be made during the classroom observations and would only be made subsequent to implementation, once the researcher had left the site. It also described how repeated reflection was made on data which is also recommended by Guba and Lincoln who suggest continual scrutiny of data. Denzin (1971, p. 177) expresses the importance of an "*intimate familiarity*" with all sources of data so that the researcher is able to sufficiently judge what is necessary to build a representative case.

An element of participant bias is also a potential threat to credibility. This occurs when a participant is aware they are part of a study and therefore attempts to "help" the researcher and behaves in such a way that is not natural but that they think is appropriate. Often, the students were not aware that the OAs had been designed by the researcher herself and she was merely an external observer. Given the performance management procedures that schools induce and training PGCE students in schools, it was believed that students are generally used to having additional persons present in their lessons, often some they are not familiar with. Initially it was also hoped that the researcher would be able to carry out multiple visits to classrooms prior to implementation to develop familiarity. Unfortunately however, this was only possible with two of the eleven groups.

A frequent and widely used technique among the literature for enhancing validity of research is triangulation. Stake (1995) emphasises the value of triangulation in improving credibility and highlights four different sources of triangulation, parallel to those proposed by Denzin (1984). Such sources are: data source, investigator, theory and methodological. This study makes use of the first, second and fourth.

Data source triangulation involves looking for consistencies in data across different contexts (Tellis, 1997) and can be done through collecting data from two sets of participants (Stake, 1995). In this study this was achieved through collecting data from both students and teachers. Although data collected from teachers was aimed to address additional areas, such as the practicality of OA implementation, much of it was also parallel to that collected from students. Responses from teachers and students could therefore be compared in order to identify commonalities and differences.

Many researchers also recommend carrying out 'member checks' (Guba and Lincoln, 1981; Hitchcock and Hughes, 1995; Merriam, 1988; Stake, 2010). This is when the researcher explains their observations and interpretations to the participants of the study to see if they approve and agree. This was partly achieved through the student focus groups and teacher interviews, subsequent to implementation. The structures of these were formulated based upon the researcher's interpretations of the data collected in the questionnaires and from classroom observations. Lincoln and Guba (1985) describe this technique as a means of assessing intentionality – that is, was it the intention of the participant to lead the researcher towards these particular interpretations and conclusions? This process of the study simply allowed participants to express whether or not, the researcher's interpretations were acceptable depictions of their views and perceptions.

This process also ties closely with the second method, investigator triangulation. This method involves having more than one investigator or researcher examining the same phenomenon (Tellis, 1997). This was utilised whereby inferences made by the researcher and 'member checks' from participants were also presented to the supervisor of this study. He had not been present for any of the

data collection so could only make judgements based on the data that was presented. As he was not directly related to the study, he was also considered to be an unbiased source.

The final method, methodological triangulation, is perhaps the most widely used, or at least widely recognised (Stake, 1995). Mixed methods research therefore, by its very nature, utilises triangulation. Quantitative and qualitative paradigms and the specific techniques within both, have their unique advantages and limitations. Often the errors and limitations yielded by one paradigm, is countered by the strengths of another. Where methodological triangulation is utilised, the research is not basing their inferences on a single source of evidence. Instead, they can triangulate multiple sources and build a more rigorous, robust case.

This particular study uses both quantitative (questionnaires) and qualitative (classroom observations, focus groups and interviews) techniques for data collection. Where something comes to light in one source of evidence, its validity is promoted where it is also evident in another source. If multiple methods produce the same evidence, we can be more confident that such evidence is valid.

8.2 Confirmability

Here the positivistic notion of construct validity, referring to whether or not the applied measures in a study are effective in addressing the concepts under investigation (Yin, 2014), is replaced with confirmability. This is particularly testing where research seeks to investigate a latent construct. As latent constructs cannot be measured directly, the measures put in place to infer such a construct must be relevant and accurate. This is a particular threat to constructivist research due to the subjective nature, where results are formed largely from interpretations by the researcher. A key feature of confirmability, as highlighted by Guba and Lincoln (1981) is that it is the researcher's responsibility to report the data and interpretations in such a way that if necessary, it can be confirmed by additional sources.

Part 1 of the PPQ was trialled and thoroughly validated through exploratory factor analysis and reliability analyses and therefore is not of concern (the process of which can be found in Bartlett *et al.* 2018). It is therefore the classroom observations, focus groups and interviews that were of greater concern and could be largely influenced by the researcher. Confirmability is largely combatted through the same techniques used to address credibility.

Triangulation of data is invaluable, and the three instruments for data collection in this study provided a spectrum of perspective into the particular scenarios within each case and could also be triangulated to confirm or refute arisen points. Yin (2014, p. 47) also recommend establishing a "*chain of evidence*". This ties in with the importance of providing detailed and descriptive accounts of processes and interpretations. A chain of evidence makes it clear to the reader how each phase of the

study follows on from the previous and how the initial research question led to resulting conclusions and is compatible with the sequential explanatory design. Given the design of this study, evidence was triangulated across instruments and a greater depth was reached. This allowed the researcher to establish a chain of evidence that was used to supplement and develop results from the earlier data collection phases. The more the findings across the individual sources and instruments were reflective and supportive of one another, the more confident we can be in that the data represents the constructs it set out to measure.

8.3 Transferability

External validity describes the extent to which a study can be generalised to people, settings and times beyond those of the given study (Campbell, 1957; Trochim, 2006; Yin, 2014). The constructivist alternative employed here is *transferability* (Reige, 2003). External validity and generalisability is, unsurprisingly, often argued to be one of the biggest limitations of case study (Zainal, 2007) and constructivist research. Generalisability is achieved when the researcher can rightly say that x will lead to y in z circumstances. However, the z is a major problem for educational research (and other social science research) as there is a myriad of uncontrollable variables in play (Bassey, 1981). Merriam (1988) also highlight that researchers who chose to follow a case study design do so not with the intention of investigating what is true for the many, but instead with the intention to gain a deep understanding of a particular scenario and what this might imply for other scenarios.

Generalisability has already been touched upon several times in terms of how it was not an appropriate target for this study and instead the alternative definition of transferability is employed. Where the traditional scientist is able to set up a controlled environment and manipulate variables at will, it is impossible to manipulate and control all influential variables in studies involving human behaviour. So many factors shape a classroom, such as gender ratio, free school meal eligibility, performance. Each student and teacher also has their life outside of the classroom that will subconsciously be brought into the classroom and influence that individual's reality. These numerous factors and characteristics are unique to each individual, each classroom and each school, it was therefore expected that noticeable variations would be evident across groups, as each event is distinctive, it cannot and *should not* be generalised to the wider population (Mack, 2010).

Therefore, as with many constructivist studies, generalisability in the traditional sense is not an objective. That said, the research does need to hold some relevance and applicability to the educational research field in order to merit its worth. Many researchers have proposed a number of alternative measures and substitutes depending on the research design that is employed.

Here, the researcher sought to provide findings that were *transferable* (Guba and Lincoln, 1981; Reige, 2003) to other settings. Instead of prescribing generalisations around what will work and what will not work across *all* classroom settings, the researcher describes what did and did not work and under what contexts and circumstances. The researcher identifies best practice in each individual environment and actively seeks differences among groups as well as commonalities. The findings of this study therefore provide a narrative of the entire process, explaining what happens and under what particular contexts. This allows teachers or practitioners to follow the classroom events and student experiences and identify for themselves how these may be complementary and transferable to their own classroom environment. It was the task of the researcher to provide data that was rich and detailed enough for the reader to determine whether the findings from the specific context are transferable to their own classroom (Yilmaz, 2013; Lincoln and Guba, 1985).

This approach is considered comparable to *naturalistic generalisations* proposed by Stake (1978, p.6). These generalisations are formed when similarities of objects or scenarios are easily recognised both in and out of the context. This requires comprehensive understanding of the particular scenario that enables the individual to recognise similarities in foreign contexts, (i.e. another classroom). This experience leads to a comprehensive awareness of why such things are as they are, where they are likely to arise and how people regard them. Bassey (1981, p.85) argues a similar standpoint and argues that an educational case study should not be judged on generalisability but upon the extent to which the details can be related both adequately and appropriately to a teacher in a similar situation, he terms this *relatability*.

Here, the researcher considers that in order to gain new insights and make new discoveries, deviations from universal laws are not only necessary but crucial. If educational research has a purpose in contributing to improving educational practice, it is not appropriate to combine case studies into one edifice (Bassey, 1981). Bassey's reasoning is interpreted to mean that improvements cannot be made without change and therefore anomalies and deviations from previous findings are necessary in order to facilitate such changes. This approach is considered comparable to miraculous medical case studies that lead to developments and understandings within the field. Here, it is not their generalisability or applicability to the broader population but their unique and unexpected characteristics that led to such advances.

This gains substantial support, for example, Kind *et al.* (2007) argue that although qualitative methods are limited by their ability to generalise, they make up for it in the depth and detail that they provide. When a case study design is chosen, it is under circumstances where the researcher seeks to gain an in-depth understanding of a particular scenario or environment, rather than to discover a general rule.

The researcher therefore aimed to enhance the transferability of the study by describing the process of the study and relevant data in as much depth and detail as possible. That way, the reader can

understand the context of where these results transpired and identify whether they are applicable and transferable to their own classroom.

8.4 Consistency

The positivistic reliability measurement is replaced here by consistency. Many issues facing reliability have already been discussed in the preceding sections that address validity. In particular, many researchers believe that internal validity is not possible without reliability (Merriam, 1988; Lincoln and Guba, 1985).

Reliability aims to minimise both the error and bias within a study (Yin, 2014). As a result, if the study was to be repeated, it would return the same results (Merriam, 1988). The difficulties in universal generalisation in educational research have been discussed where each classroom is distinct and possesses characteristics that deviate from the next. Though the investigation can be repeated in another environment, it is not possible to replicate the scenario or environment. Instead of producing results that can be reliably transferred to other contexts, the aim of the researcher is to demonstrate to the reader that, given the data, the results and the conclusions make sense and how such conclusions were reached is clear and intuitive. Lincoln and Guba (1985, p. 288) describe this *consistency* or *dependability* of research as more applicable here.

Again we are brought back to the importance of establishing a chain of evidence and following a process of external verification. The data collection tells a story that the reader can follow and understand. It was anticipated that the findings of this study would likely reveal a range in results not only from case-to-case but from group-to-group. Although implementation of the educational OAs turned out to be highly successful in one classroom were less successful in others. However, this is not considered to be a flaw in the reliability, instead, the approach of an in-depth case study allowed the researcher to identify the various characteristics across these environments that could account for such differences. For example, x worked in classroom y under z circumstances but not under a circumstances. This brings us back again to transferability.

Merriam (1988) also draws attention to the fact that repeating a qualitative case study will also not yield the same results due to the subjectivity and interpretation of the researcher. This ties back to the theory of reflexivity and the importance of acknowledging the predispositions of the researcher. It should be emphasised that this does not discredit the original study, nor the following replication study, but simply highlights that alternative interpretations can be made on the same data. So long as any biases are acknowledged and the researcher is descriptive in their reasoning and justification for such interpretations, the reader can judge whether they agree.

However, the process of peer debriefing is put in place in order to ensure such interpretations are rational and justified. Although interpretations will vary from person-to-person, each of these persons should be able to appreciate those of others and understand their thinking behind such inferences. Where inferences are made that cannot be understood by external individuals and appear to be an entirely biased output of the researcher, the study cannot be considered as consistent.

8.5 Summary

This chapter has argued the constructivist alternatives for validity and reliability and why such positivist approaches cannot be applied to processes where subjective, qualitative methods are dominant.

The measures that were taken in order to enhance the validity of this study and reduce such threats have been discussed at length, however there were three overarching techniques that were considered:

- Triangulation of data collection instruments •
- Predispositions and biases of the researcher are set out initially prior to data collection
- The research process and the justifications of inferences is highly descriptive and narrated •

Table 11 summarises the techniques put in place to combat validity and reliability issues in this study.

Table 11 – Summary of Techniques to Improve Validity and Reliability					
Validity Aspect	Description	Technique for Enhancement			
Credibility	The intervention caused the effects that are seen and described	 Descriptions of everything that is witnessed during data collection Continual scrutiny of data Member checks Predispositions and biases of the researcher clearly set out Triangulation 			
Confirmability	Data collection tools measure what they intend to measure	TriangulationEstablished chain of evidenceMember checks			
Transferability	Findings can be transferred into other settings	• Descriptions of scenarios, participants and contexts			
Consistency	Conclusions are sound based on the evidence	Established a chain of evidencePredispositions and biases of the researcher clearly set out			

Chapter 9

Preliminary Data Analysis

As this study applied a mixed method approach, the researcher deemed it appropriate to first perform a surface level analysis of the quantitative and qualitative data separately before full, synthesised analysis was carried out. This was for two reasons:

1/. As this study was a sequential explanatory design whereby one method of data collection guides another, it was important to identify what the PPQ and classroom observation data revealed so that the follow-up focus groups could be designed in order to gather deeper and relevant information.

2/. It was considered important to recognise any patterns and themes among the quantitative and qualitative data in order to identify their applicability to the research questions and to identify a suitable evaluation framework for full analysis.

This chapter therefore provides a surface level breakdown of the results from the Perceptions of Physics Questionnaire and the Classroom Observations from both case studies.

9.1 Quantitative Data Analysis

Quantitative analysis was carried out on all responses gathered in Part 1 of the PPQ, the Likert-scale. An independent sample t-test was computed using SPSS version 23 in order to explore whether students showed any differences in their perceptions of their day-to-day physics lessons (pre-PPQ) and of the implementation activity (post-PPQ).

Before data analysis on the PPQ could be carried out, successful validation of the instrument had to be confirmed, the process of this is detailed in Bartlett *et al.* (2018). Factor analysis revealed six reliable and valid scales with the pre-PPQ. These were Future Aspirations in Physics, Interest in Astronomy and Space Science, Confidence, Perceived Relevance, General Interest and Investigation Work.

Students responded to a number of items within each of these scales on a 5-point Likert Scale where 1 = Strongly Disagree, 2 = Disagree, 3 = Neither, 4 = Agree and 5 = Strongly Agree. Items were generally positively phrased (e.g. I enjoy my physics lessons) and those that were negatively phrased (e.g. physics lessons are boring) were reverse coded for analysis. This meant that high scores indicated a positive opinion and low scores indicated a negative opinion across all items.

Before data is separated into cases, Down2Earth and Hubble, study-wide implementation of the OAs was analysed in terms of this study's overall influence on students' learning experiences.

Independent sample t-tests were computed in order to compare students' PPQ scores across the perception constructs on occasions of pre and post implementation of the OAs. Table 12 summarises such scores across the total number of student responses (N) in the pre- and post-test in terms of their mean and standard deviation (SD) and the significance value (p) and effect size (d) of the t-test. Effect sizes are provided both as values and size (small, medium or large) according to Cohen (1988). Scales highlighted in green are indicative of significant differences between occasions. Where the p value (significance) is equal to 0.05 or less, we can be 95% confident that such results are not due to chance. Alternatively, the effect size describes the magnitude of this result: small, medium or large.

It is observed that the scale regarding students' Future Aspirations to Physics is not including in the following reports. This is because this scale was only measured in the pre-PPQ. It was not measured in the post-PPQ as the researcher did not anticipate that a change toward this construct was possible in a one-lesson initiative. Students' initial perceptions towards future involvement in physics were collected in order to identify whether results were reflective of other existing and previous studies.

Table 12 - Pre- and Post-PPQ Scale Scores for all Students across both Cases						
Factor	N pre/post	Pre Mean (SD)	Post Mean (SD)	р	Cohen's d	
Interest in Astronomy	230/223	3.35 (0.92)	3.43 (0.86)	0.334	-	
Investigation Work	208/194	3.57 (0.77)	3.80 (0.68)	0.002	0.3 (small)	
Confidence	205/193	3.33 (0.83)	3.80 (0.71)	< 0.001	0.6 (medium)	
Perceived Relevance	230/209	3.16 (0.68)	3.26 (0.74)	0.176	-	
General Interest	233/226	3.63 (0.77)	3.89 (0.72)	< 0.001	0.3 (small)	

Quantitative results from the study are broken down into the relevant cases in the following sections. Here a summary of the results from statistical analysis are provided that is later fed into the evaluation framework and triangulated with qualitative data for full analysis and interpretation.

9.1.1 Case Study 1: Down2Earth

Results from the independent sample t-tests of students participating in the Down2Earth case study are summarised in Table 13. Overall, students provided more positive responses to each factor on the post-PPQ than on the pre-PPQ. However it is seen in Table 13 that students only reported significantly more positive responses on three of the five factors on the post-PPQ, relating to the initiative activity. Students reported significantly more positive perceptions of their Confidence (t (191) = -5.250, p <0.001), their Perceived Relevance (t (237) = -2.853, p = 0.005) and their General
Interest (t (249) = -3.298, p = 0.001) towards the activity than towards their day-to-day physics lessons.

Students' interest in Astronomy and Space Science although was seen to be slightly more positive on the post-PPQ, showed no significant difference, and the same for students' perceptions of Investigation Work.

It is considered that a one-off initiative lesson is not sufficient to influence students overall opinion regarding Astronomy and Space Science substantially. In terms of investigation work, Ofsted (2013) report high levels of investigation and inquiry in key stage 3 science which is likely why this saw no significant increase.

Table 13 - Pre- and Post-PPQ Scale Scores for Students in the Down2Earth Case Study					
Factor	N pre/post	Pre Mean (SD)	Post Mean (SD)	р	Cohen's d
Interest in Astronomy	125/120	3.18 (0.94)	3.30 (0.87)	0.346	-
Investigation Work	102/91	3.75 (0.68)	3.83 (0.65)	0.406	-
Confidence	103/90	3.30 (0.87)	3.92 (0.77)	< 0.001	0.8 (large)
Perceived Relevance	125/114	3.09 (0.71)	3.37 (0.75)	0.005	0.4 (small)
General Interest	129/122	3.53 (0.80)	3.85 (0.74)	0.001	0.4 (small)

Figure 7a and 7b illustrate the distributions of student scores on each of the five scales on the pre-PPQ and on the post-PPQ within the D2E case study.



Figure 7a – Distributions of D2E Case Study Student Scores on Individual PPQ Scales on Pre-test



Figure 7b – Distributions of D2E Case Study Student Scores on Individual PPQ Scales on Post-test

Figures 7a and 7b are illustrative of the skewness of the data on both occasions. Skewness describes the symmetry of a statistical distribution by assigning a value to the extent at which a distribution differs from normal. A negative skewness value is demonstrative of the mode of the data being of greater value than the mode. A positive skew is demonstrative of the mode being of smaller value than the mean. Table 14 details the specific skewness values for each of the scales in the pre- and post-PPQ.

All of the scales on the pre-PPQ possess a skewness value within the range of -1 and 1 which was used in this study as the normal distribution range. Investigation Work shows the greatest skewness which is negative, thus demonstrating students had a slightly more positive perception of these items.

However, Figure 7b suggests much greater skewness in the scales on the post-PPQ and this is reflected in the corresponding values. On the post-PPQ students in the Down2Earth case study yielded scores outside of the normal distribution range for factors Investigation Work, Confidence and General Interest. All of these factors were negatively skewed and thus demonstrate that on the post-PPQ, students' perceptions were more positive in these areas. Nonetheless, such values still remain within the range of -2 and 2 which is considered a normal distribution by many researchers (e.g. George and Mallery, 2010).

Table 14 – Skewness Values of D2E Student Scores on Pre- and Post-PPQ				
Fastor	Skewness			
ractor	Pre-PPQ	Post-PPQ		
Interest in Astronomy	-0.289	-0.359		
Investigation Work	-0.937	-1.137		
Confidence	-0.824	-1.272		
Perceived Relevance	-0.087	-0.467		
General Interest	-0.777	-1.511		
Future Aspirations	-0.086			

9.1.2 Case Study 2: Hubble's Expansion of the Universe

-

Analysis of responses from students in the Hubble case study revealed slightly different results. Here, generally students reported higher scores on the post-PPQ for all factors apart from Perceived Relevance (see Table 15). Students therefore indicated that they found their day-to-day physics lessons more relevant than the initiative activity, though this was not significant.

In this case just two factors showed significant differences in the pre- and post-PPQ, namely Investigation Work (t (207) = -3.557, p <0.001) and Confidence (t (203) = -3.321, p = 0.001). In key stage 5 the curriculum places less of a focus on skills compared to lower key stages (Ofsted, 2011; 2013), and subsequently fewer practical lessons are seen to take place. Therefore it is not entirely

Table 15 - Pre- and Post-PPQ Scale Scores for Students in the Hubble Case Study					
Factor	N pre/post	Pre Mean (SD)	Post Mean (SD)	р	Cohen's d
Interest in Astronomy	105/103	3.54 (0.86)	3.59 (0.82)	0.710	-
Investigation Work	106/103	3.39 (0.82)	3.77 (0.71)	< 0.001	0.5 (medium)
Confidence	102/103	3.37 (0.79)	3.70 (0.63)	0.001	0.6 (medium)
Perceived Relevance	105/95	3.24 (0.63)	3.12 (0.70)	0.199	-
General Interest	104/104	3.74 (0.71)	3.92 (0.69)	0.063	-

surprising that in the Hubble case study, a significantly higher score was obtained on the post-PPQ in relation to this scale.

It is disheartening to see that even students in the Hubble case who have chosen to pursue physics beyond the compulsory level; and report a generally positive general interest (3.74), still have relatively negative perceptions of their ability in the subject and the relevance of physics.

Figure 8a and 8b illustrate the distributions of student scores on each of the five scales on the pre-PPQ and on the post-PPQ within the Hubble case study.

As with the D2E case results, all scales in the PPQ demonstrated values within the normal distribution range of -1 and 1 on the pre-occasion (Figure 8a). However, a normal distribution was also seen on the post-PPQ scores as illustrated in Figure 8b and in Table 16. This implies that students in the Hubble case study reported less 'extreme' perceptions towards their learning experiences than those in the D2E case study.



Figure 8a – Distributions of Hubble Case Study Student Scores on Individual PPQ Scales on Pre-test



Figure 8b - Distributions of Hubble Case Study Student Scores on Individual PPQ Scales on Post-test

		e		
Factor	Skewness			
ractor	Pre-PPQ	Post-PPQ		
Interest in Astronomy	-0.564	-0.585		
Investigation Work	-0.432	-0.424		
Confidence	-0.579	-0.129		
Perceived Relevance	-0.214	-0.003		
General Interest	-0.472	-0.995		
Future Aspirations	-0.131			

Table 16 – Skewness Values of Hubble Student Scores on Pre- and Post-PPQ

9.2 Qualitative Data Analysis

Preliminary qualitative analysis was only carried out by means of thematic analysis (TA) on classroom observation data and the open, qualitative questions included in the PPQ. Thematic analysis is a process used to identify, analyse and find patterns within a data that are relevant to the subject phenomenon (Joffe, 2012; Braun and Clarke, 2006). However, there is no unanimous agreement among researchers on specifically how the analysis should be applied (Nowell *et al.*, 2017; Braun and Clarke, 2006). Results from TA combined with the quantitative analysis is what was used to inform the semi-structured student focus groups and teacher interviews.

Given the inductive nature of this study, themes were generated from the raw data rather than deductively based on an existing theory (Nowell et al., 2017) and was therefore data-driven rather than theory-driven (Braun and Clarke, 2006).

Nowell *et al.* (2017) highlight how inductive approaches can yield themes that appear irrelevant to the particular research questions of a study. However, this is another purpose of the evaluation framework that discussed in Chapter 5 and applied later in order to synthesise all quantitative and qualitative data and prioritise that of importance and relevance.

Given that this process focussed on broad learning experiences that occur in the classroom to guide the overarching evaluation framework, data was not separated according to case or group and was instead combined for this stage of analysis.

Figure 9 illustrates the full process of analysis. This is comparable to Figure 6 in sub-chapter 7.3.1 that illustrates how observation notes were revisited and reviewed on multiple occasions subsequent to the actual observation in a reflexive process (Cohen *et al.*, 2011). Once the final reflection had been made for each observation across both case studies, TA was carried out.



9.2.1 Thematic Analysis of Observational Data

All observation data was colour coordinated according to group, and each observation and each sentence was isolated and printed out. The researcher then laid out each of these observations and began to review them, coding them into numerous categories and then combining these into broader themes and patterns among the data. This manual thematic analysis was perceived to be preferable over the use of software, as it allowed her to view all the data at once making it easier to connect data and identify relationships.

Where observation data was revisited, it was isolated by group. However, the thematic analysis involved combining data across groups and cases. Consequently, this allowed for further reflexivity and interpretation of scenarios, and the identification of behaviours, communication and implementation styles and interactions not only across students within groups but also between groups.

Similar to observational records, the thematic analysis was revisited several times. Upon identifying all apparent themes in the observations, the researcher recorded these themes and the relevant observations and would then start the thematic process again a few days or a week later. This was done so that the researcher could review all the data again and explore any themes that were not

identified previously. Following this, the researcher turned to the literature to explore qualitative research around classroom behaviour, communication and engagement in order to identify other circumstances and scenarios where similar observations had been made and the interpretations drawn.

Table 17 summarises the six final themes considered to be present among the observation data. All observations considered to be applicable to each of these themes was so classified and also categorised to whether this was a positive or negative observation. Specific examples of observations considered to be applicable to each of these themes is provided in Appendix 13.

Table 17 – Thematic Analysis of Observational Data		
Theme	Explanation	
Confidence	Evidence of student demonstrating confidence (or lack of) in their ability.	
Engagement	Evidence to suggest students were (or were not) engaged with activity	
Processes of Inquiry	Evidence of students carrying out additional elements of inquiry, beyond scope of activity	
Student-Teacher Interaction	Evidence of the types of relationship between the teacher and students, both academically and socially	
Situational Interest	Evidence to suggest students were (or were not) interested in the activity or context	
Student Relationships	Evidence of the types of relationships between students.	

9.2.2 Content Analysis of PPQ Open Questions

Content analysis on the PPQ was carried out on two open questions in the pre-PPQ and three in the post. Content analysis is different from thematic analysis in that it primarily categorises data rather than interprets the themes among it.

The two pre-PPQ questions asked students what they liked most and least about their day-to-day physics lessons, and two of the post-PPQ questions were parallel to these but in relation to the implemented activity. The third question in the post-PPQ asked students what they thought could be improved in the activity. Answers on the post-PPQ were also separated by case as it was evident that many responses were specific to the OA.

Here, the content analysis was much more straight-forward than the thematic analysis of classroom observations as there was much less data to work with. Table 18 summarises the themes identified among students' responses, examples of each theme can be found in Appendix 14.

Table 18 – Content Analysis of Open PPQ Responses		
Question	Themes	
	Practical lessons	
	• Learning	
	Astronomy	
What do you like most about physics lossons?	• The teacher	
what do you like <i>most</i> about physics lessons?	Mathematics	
	Particle physics	
	• Diversity	
	Interesting	
	Writing	
	Boring	
	• Difficulty	
	Specific topics	
What do you like <i>least</i> about physics lessons?	Practical lessons	
	Reading	
	Working with graphs	
	Mathematics	
	• The teacher	
	• The simulator	
	Investigation	
What did you like <i>most</i> about the activity?	Asteroids and craters	
(Down2Earth)	• Independence	
	• IT element	
	Blowing things up	
	• Learning	
	Writing	
	Restrictions	
What did you like <i>least</i> about the activity?	• Repetitive	
(Down2Earth)	Investigation	
	• Difficulty	
	Distractions	
	• Time	
	Additional tasks	
	• Investigation	
How do you think the activity could be improved?	Group work	
(Down2Earth)	• The simulator	
	• Writing	
	• Guidance	
	• Freedom	
	• Other	
	• The age of the Universe	
	• Graph plotting	
What did you like <i>most</i> about the activity?	• Working with real data	
(Hubble)	• Finding Hubble's constant	
	Group work	
	Independence	
	Independence	
	Graph plotting	
	• Kepetitive	
what did you like <i>least</i> about the activity?	• 11 elements	
(הונטטופ)	Iviatnematical aspects Calculators	
	Calculators Other	
How do you think the activity could be improved?	• Ouler	
(Hubble)	• IT elements	

•	Time
•	Interactivity
•	Instructions/guidance
•	Other

9.3 Summary

This chapter has detailed the initial, separate procedures of data analysis that was carried out for quantitative data and qualitative data. Though this is a mixed method study that makes use of a robust evaluation framework for analysis, the researcher deemed it important to initially review collected data according to their separate paradigms. Analysis of the quantitative data provides an overall picture of what took place in the study and students' perceptions of their learning experiences in day-to-day physics lessons and in the initiative lesson. Classroom observation data demonstrates some of the underlying themes and patterns among the classrooms that were not necessarily addressed in the closed questions in the PPQ. Preliminary thematic analysis, particularly of observational data demonstrated the substantial amount of data that was yielded in the researcher's inductive approach, as anticipated. However, in order to enhance the transferability of such data, the reader must understand the contextual elements of each implementation group. The evaluation framework therefore not only synthesises all the data that has been collected but also ties such evidence with its relevant context in order to better understand variations across students' learning experiences. All data is therefore fed into the evaluation framework in order to refine evidences and address the research questions in detail. This is carried out in Chapters 10 and 11.

Chapter 10

Case Study 1: Down2Earth

10.1 Goal Orientation

The OAs that formed the case studies were written approximately one year prior to implementation in this study. At this point, the researcher had very little pedagogical knowledge, resulting merely from interactions with mostly European teachers at teacher training events.

As the OAs follow processes of inquiry, the researcher considered that the research question or scenario that was presented to students for investigation stood as their lesson objective. If students were able to complete an investigation that would lead them to answer this question by the end of the lesson, then it was considered that they had achieved the objective. There was therefore no strict mark or grading scheme. However, through implementation and development of the researcher's pedagogical and educational understanding, it unfolded that this was not specific enough for the day-to-day classroom. Quinn (1996) argues that objectives should be unambiguous and students should not have to assume the lesson objectives or guess what it is that is expected of them.

Passing feedback was received from the teacher at group Brown during implementation:

From a pedagogical point of view, they [the students] would benefit from WALTs and WILFs, so clear learning objectives set at the start and ways to measure them at the end.

WALT is an acronym for $\underline{W}e \underline{A}re \underline{L}earning \underline{T}o...$ thus outlining the learning objectives of the lesson set out at the beginning and WILF is an acronym for $\underline{W}hat \underline{I}m \underline{L}ooking \underline{F}or...$ that is reflected on at the end of the lesson and describes what the students are looking for in terms of success criteria to demonstrate that they have achieved the objectives.

The researcher also reflected through increased experience and exposure to classrooms that there should not be a single, overarching objective and instead had there been a more constructive breakdown of smaller objectives, this could have also aided the teacher in their implementation and differentiation. That is, the teacher may have been more readily able to adjust implementation style in order to meet the needs of the different students in their class if multiple, progressive objectives had been laid out.

Despite these arguments, results from the PPQ suggested that this was not largely detrimental to students' learning experiences. One particular item on the PPQ was 'I learn physics quickly' (pre) and

'I learnt how to do the activity quickly' (post). It was assumed that had students struggled with understanding what they needed to do in the OA lesson, they would have scored low on these items. However, this was not necessarily the case.

Students in this case study responded significantly more positively on the post-PPQ (mean = 3.83 ± 0.93) than on the pre-PPQ (mean = 3.16 ± 1.06), t (196) = -4.664, p < 0.001. This result indicates that students felt they learnt what they had to do in the OA, more readily than in their day-to-day physics lessons.

This item was also one of the items that made up the Confidence scale, which as highlighted in Chapter 9, showed extremely promising results with regards to the D2E activity. Students revealed significantly higher confidence on the post-PPQ (mean = 3.92 ± 0.77) than on the pre PPQ (mean = 3.30 ± 0.87), t (191) = -5.250, p < 0.001, therefore demonstrating students held higher confidence towards the activity than their day-to-day physics lessons.

Figure 10 illustrates the responses to these items on the post-PPQ of all students in the D2E case study. Here it can be seen that the most common response for all of these statements was Agree. The next most popular response was Strongly Agree for five out of the seven statements, while *I think I will get good marks in the activity* and *the activity was useful to my learning* revealed Neither as the second most common response.



Figure 10 – Student Responses to Confidence Scale on Post-PPQ in the D2E Case Study

Though the researcher does not dispute that the objectives of the activity should be set out more clearly and more differentiated, it is still heartening that this did not appear to impede on the majority of students' experiences.

Nonetheless, although most participants indicated a level of agreement, it is also clear from Figure 10 that there were some that did not. Table 19 summarises the percentage of students in the D2E case study that either disagreed or strongly disagreed with the items in the Confidence scale. It is possible that these students would have held higher perceptions of confidence had the objectives been clearer and they had a better idea of what was expected of them and what they should have been aiming to achieve.

Table 19 – Respondents who Disagreed or Strongly Disagreed with Confidence Scale Items		
Item	Percentage of Responses	
I was good at doing the activity	7.4	
I think I will get good marks in the activity	12.9	
I learnt how to do the activity quickly 10.0		
I understood what we learnt in the activity 6.		
The activity was useful to my learning 6.		
I found the activity difficult (recoded) 9.9		
I struggled with the activity (recoded)	8.7	

In terms of students' internal orientations; whether they were performance or learning oriented, this appeared to be influenced by the way students approached the activity. Group Pink stood out in observations as a group who were working primarily individually. Although students interacted with one another and were keen to show each other their work, their behaviour was characteristically more competitive than cooperative or collaborative. For example, during the Recreate a Crater activity, students were seen to go over to another peer's desk to see how close they'd got to recreating an existing crater, to then return to their seats and attempt to beat it. Alternatively, in other groups within this case, there were much more recordings of students asking one another questions such as how they had gotten a particular result and what parameters they had changed. This behaviour was therefore considered to be more cooperative.

Such findings led the researcher to consider that students' approach to lesson tasks, whether cooperative or competitive is a strong indicator for their goal orientations. Students who are competitive will typically possess a performance orientation, they don't seek to learn from their peers, they simply seek to beat or outperform them. Students who are cooperative or collaborative however, appear to be more supportive and learn from their peers and use them for guidance to reach a common

goal. Elements of cooperative and collaborative learning observed in the D2E case are discussed in greater depth in Chapter 10.4.

One final observation that cannot be ignored was suggestive that one particular student's goal orientation actually altered after completing the D2E activity. This student was Lucy in case Blue. In the first standard physics lesson observation, Lucy and her classmate Jess were noticeably disengaged, were talking throughout the lesson about unrelated topics and did not complete the lesson work. However, in the main implementation lesson, the two students were the first to complete the investigation and follow-up questions. Upon realising this, Lucy turned to Jess and said "*so I can do science*". This single, entirely unprompted comment demonstrated the effect just one successful lesson can have on a student's morale and confidence. It is also suggestive that before implementation, Lucy was of the belief that she could not do science which was likely based on a performance-orientated disposition and had resulted in disengagement and a disregard for science. However, where she applied an increased amount of effort in the implementation lesson and focused on the set task, she openly expressed her realisation of not being limited by her abilities, thus suggesting a shift to a learning orientation.

10.2 Differentiation

Justification for the methodology and approach of this research study has been largely based around the intricate and complex environment that is classroom learning. It has been heavily argued in this study and by other authors that generalisation in education is not feasible. Each student, teacher, classroom and school is unique and something that can be confidently said about one, cannot necessarily be said about another. Classrooms are therefore considered to be circumstantial and situational.

An educational study can therefore not be completed without due regard to differentiation. Classroom differentiation provides a diverse range of learning options that accommodate for alternative levels of ability, interest, engagement and learning styles (Tomlinson, 1995). Everything therefore, that is unique to each classroom environment creates difficulties when transferring outcomes to other classrooms.

Here it is considered that the dimension of differentiation ties closely with the previous dimension explaining the importance of goal orientation. Clear objectives help the student to understand exactly what is expected from them and what they should aim to achieve. Where such objectives develop from one another and follow a progressive process, they also aid in differentiation of the OA.

Differentiation is considered by the researcher to be particularly difficult to account for by an individual who is external to the classroom. Where teachers are familiar with their class and know

their students' individual strengths and weaknesses, these are entirely unknown to the external OA developer, in this case, the researcher. This was echoed by a physics PGCE student who was reviewing the D2E activities. She explained:

I could run this activity with the class at the last school I was in and it would fit into a single lesson nicely, but if I tried to do it with the class I have at the school I'm in now, there's no way they'd be able to complete it in one lesson.

All of the groups participating in the D2E case study were set classes, i.e. they were groups of students considered to be of similar ability. However there still remains a spectrum of ability and learning needs within such classes and Ofsted (2013) emphasise that class setting alone cannot be relied on for complete differentiation.

Despite this, when the researcher met with teachers ahead of implementation to go through the activities and discuss various approaches to implementation, in no classroom were any of the D2E activities adjusted or implemented differently for different students, including those implemented by the researcher.

Though teachers across groups had varied implementation styles and provided different levels of structure and guidance, this was their method of differentiation for the entire class. All students in the same group experienced the same implementation and there were no differences between students.

For example, Mr Navy had removed some of the investigation follow-up questions from the Recreate a Crater activity, as he felt that his students would not be able to get that far in a single lesson. Mr Blue also implemented the Making a Meteorite over two lessons, again basing this on the work pace of his students. However, no methods of differentiation within the activity were embedded for individual students and so the influence of this on students' learning experiences is explored.

Differentiation is first addressed through results from students' confidence in their day-to-day physics lessons and in the D2E activity as measured in PPQ responses and recorded in classroom observations. Though it is hard to determine an optimum level of confidence, that is where the lowest score is 1.00 and the highest score is 5.00, a score of 5.00 would indicate the highest possible confidence but may also imply that the student is not being challenged enough in their learning. However, a score of 1.00 would imply that the student was severely struggling with the lesson content and therefore risks students' sense of competence which according to SDT is essential for engagement and motivation. It is therefore at the discretion of the researcher that a score of 4.00 was optimal.

Chapter 9 has already demonstrated the significantly more positive scores reported by students on the Confidence scale on the post-PPQ. Figure 11 illustrates the distribution of scores on the Confidence scale on the pre-PPQ (left) and post-PPQ (right) for students in the D2E case study. In the pre-PPQ, when students were referring to their day-to-day physics lessons, results yielded a mean score of 3.30

 (± 0.87) , indicating only slightly positive scores past an indifferent score of 3.00. The most common average score to items on this scale was 3.40, therefore not reaching the firm positive score of 4.00.

Figure 11 – Histograms of the Distributions of Student Scores on Confidence Scale on Pre (left) and Post (right) D2E Implementation



For the post-PPQ, when students were referring to the activity, the same scale yielded a mean score of $3.92 (\pm 0.77)$ indicating a more positive score and a skewness in the distribution of -1.272, therefore outside the normal range of -1 to 1. Although a mean score of 3.92 could be considered as not a very promising result as it would in any of the other PPQ scales, as discussed, the optimum score in this instance was considered to be 4.00. A score of 4.00 was also the most common score obtained by students in the case-study and 64% of students yielded a score of 4.00 or higher, therefore demonstrating that the majority of students in this case study felt confident in completing the activity. This is compared with the 27% of students who reported a score of 4.00 or higher in the pre-PPQ.

It is equally important for differentiation to aid lower ability students and prevent them from struggle or failure as it is to ensure that higher ability students are challenged and do not become bored of classroom tasks. Consequently, it could be argued that students who reported a score of 5.00 (strongly agree) to the Confidence scale were not challenged enough in the activity and therefore run a risk of disengagement (Ofsted, 2013), however this was only reported by eight students (8.9%).

Alternatively, 13 students (10%) reported a score of less than 3.00, indicating that these students struggled with the activity and would have benefited from additional guidance and support.

The post-PPQ also asked students to report what they liked most and least about the activity. Here, the variations in students' experiences of the activity as indicated by responses to the Confidence scale, is further evidenced in such responses. The following statements demonstrate that some students experienced difficulties in the activity and others found it simple.

What did you like least about the activity?:

"The kinetic energy, I don't understand it"	(Female, group Green)
"It was a bit complicated at times"	(Male, group Navy)
"The middle part where I couldn't figure out the number"	(Female, group Brown)
"It was hard"	(Male, group Brown)
"How hard it was"	(Female, group Brown)
"We didn't have to think too much"	(2x responses, Male, group Blue)

The parallel responses from students in group Blue imply that they found the activity too straight forward and it didn't challenge them. However the remaining students' responses indicate that they had difficulties with the activity. Interestingly there were no conflicting responses regarding level of difficulty reported by students in the same group.

Reports of difficulty from students in group Brown were more expected due to the younger age of participants. As discussed, the D2E OAs were designed for year 9 students whereas group Brown was a top set year 8 group.

In an interview with the teacher of group Brown, the researcher asked about the applicability of the OA to students in the corresponding group.

Researcher:	Can you describe the type of students that you think would not benefit from this activity?
	ucuvny.
Mr Brown:	Yeah it would be those that can't follow a set of written instructions, of which there
	are, there are a few in that class. At least I can definitely think of a couple of those
	that have um autism within the class, where the lack of structure would have
	been a bit of a barrier for them.

With regard to this comment, it is possible that the four students who responded with the struggles detailed previously in their post-PPQ responses have additional learning needs and therefore the activity did not accommodate for this smaller group of students within the class. This emphasises the importance of differentiating activities not only on ability but also on additional learning needs or preferences.

Responses from students in group Brown who participated in the focus group however, did not imply any problems with difficulties in the activity and if anything, implied the activity process was clearer than some of their day-to-day lessons.

Researcher:	Can you name three good things and three bad things about your normal physics lessons?
Jack:	The teacher makes it obvious what we need to do
	[pause]
Researcher:	It doesn't have to be three, just any good things or bad things you can think of.
Jack:	Sometimes the work that we do can be stretched a bit further than we need to go so it becomes quite confusing sometimes
Jack:	Sometimes it's not a problem because I like learning new stuff but sometimes it can be confusing
Isabel:	Yeah sometimes the majority of people don't get the subject and then it goes on further

Later on in the focus group, when students were asked the same question in relation to the Down2Earth activity, the only bad things that were voiced by students were computer-related problems. One female student, Isabel, also reported a positive of the activity being that "*It was good because the instructions were really clear to follow, like they were clearly laid out*".

Group Brown was the only group where the researcher was able to collect the student worksheets back. As students recorded all of their work during the activity on these worksheets, they could be used to review how the students got on. This revealed that 14 students completed the entire worksheet (8 M, 6 F), and 11 did not complete it (6 M, 5 F). Generally those who did not complete the activity got as far as recording their observations but did not answer the questions that asked students to reflect on their findings and the investigation process. Though it cannot be assumed that these 11 students did not complete the activity, purely because it was too difficult, it could be that they were distracted or spent too much time on the Impact Calculator. However this again reinforces that differentiation does not solely target ability and this is evidence that some students would have benefited from more embedded differentiation that focused on their learning style, for example, through additional structure and guidance for time management.

The next step was therefore to ask Mr Brown how this aspect could be improved in the future and from his point of view, how an OA might be best structured.

Researcher: Before implementation, all teachers are told how the activity can be differentiated. Would you rather be provided with the most involved version, the hardest so that you could add in scaffolds, or the simplest version?

Mr Brown: So would I rather start at the top or bottom?

- Researcher: Yes, so would you rather I gave you a OA with the minimum guidance and you could add in various...
- Mr Brown: Yeah I know what you're saying, um... [long pause] I would say, probably... Ooh this is one I need to think about. [very long pause]. Something that requires students to work so independently is probably geared towards your like, set one, set two pupils...would be my feeling. So yeah, probably the higher end and differentiate down. But that's not to say... that's not to say that it would be too difficult to make the higher end and turn it into... if you'd gone to the trouble to make the higher end, it wouldn't be too difficult to take some information out and make middle and lower OAs. So you had high, middle and lower, I think that would probably be easier.

Prior to any implementation, the researcher had assumed that it would be best to provide teachers with everything and they could then remove scaffolds and guidance where necessary in order to assign more of the work process to the students. However, Mr Brown's feedback implies that this is not necessarily the best way to equip teachers. Although the researcher had initially anticipated that this would leave less work to the teacher, it transpired that in fact, even the process of editing the OAs slightly is not preferable. Instead, Mr Brown implied that teachers would want to be able to deliver the OA right off the shelf.

Generally, the researcher considers the results of students' confidence and the majority of classroom behaviour observed among groups in the D2E groups as a success. Given that a point commonly raised in the literature describes student perception of their ability as a highly influential factor in their liking of the subject (Barmby and Defty, 2006; Barmby *et al.*, 2008; Williams *et al.*, 2003) and that common reasons for not pursuing physics (or science in general) post-compulsory level is because it is too hard (Gill and Bell, 2013; Tripney *et al.*, 2010) even more concerning is that this has also been reported from students who would make strong maths and physics candidates (Rodd *et al.*, 2014). It is therefore essential to encourage and nurture students' confidence in physics if we are to inspire post-compulsory physics education and physics-based careers.

However, the importance of challenging students in lessons is also recognised and that if activities are not differentiated in order to benefit the higher ability students within the class, then this also runs the risk of disengagement (Ofsted 2013).

On reflection, although the researcher remains of the opinion that although there is much potential and flexibility in the activity to include differentiation, for clarity and for ease of the teacher it would be advantageous to present these different versions as hard copies rather than hypotheticals.

Moving forward, OAs should and will be provided to teachers in three versions, a top, middle and bottom, rather than the researcher simply describing how the activity can be differentiated. Tomlinson (1995) emphasises that there is no single learning template for classrooms. With differentiated versions of the same activity, the teacher can more readily identify which is the most suitable for their individual students and can better accommodate for different learning styles and preferences. As all versions of the activity will still follow the same core process, and have a common goal between them, it will not prevent students completing different versions from working collaboratively.

10.3 Student Autonomy

One of Ofsted's (2011) key findings in their Successful Science report, was that schools that are seen to display the highest standard of science teaching held scientific inquiry at the core. Students were given opportunities to ask questions, design and carry out investigations for themselves, therefore providing them with a greater degree of control over their learning. Autonomous learning is arguably one of the most imperative characteristics of constructivism and IBL as they focus on a student-centred approach to learning. Consequently, if students' did not feel a sense of control over their learning during the activity, this threatens the inquiry-based nature of the OAs.

Given the investigative style of the Down2Earth activities and the use of an online simulator as the "lab" or "experiment", it was anticipated that they would provide an opportunity for students to work more autonomously than they would in a traditional lesson. Where even in practical lessons, students typically have to share equipment with their peers, provided that schools had access to enough computers for each student (which was the case in all groups), students could manipulate the simulator entirely of their own initiative. This is not to say students could not still work collaboratively, but that they were also able to take control of their learning and work through the activity at their own pace.

Evidence for the D2E activities enabling students to work autonomously was provided in focus group feedback from students in group Brown. Early on in the focus group students were asked to reflect on who they felt was in control of their learning in their physics lessons.

Researcher:	Who would you say is in control of your learning in physics lessons?
Jack:	Definitely the teacher
Multiple Students:	Yeah [in agreement]

Researcher:	Yep, and what makes you say that?
Claire:	Because he knows what to do
Isabel:	Yeah and he has the knowledge

However, later on in the focus group (seven questions later), students were asked a parallel question but in relation to the activity.

Researcher:	<i>Who was in control of your learning during the activity?</i>
Isabel:	Oh, us! Because we chose the options to what like, rocks we wanted to use
Claire and Jack:	Us
Jack:	[following on from Isabel] yeah and what size meteor

These two contrasting responses are palpable evidence of the shift in students' perceptions of who was in control of their learning in day-to-day lessons compared to the OA. Pre-implementation observations of group Brown took place in two lessons that both involved practical activities. However, on both occasions the format of the practicals involved students gathering around the teacher who demonstrated the process, students would then break off into groups to replicate the experiment and record their measurements. In the implementation activity however, the only demonstration that the teacher provided was a brief introduction to the Impact Calculator and how to adjust the parameters, the rest was down to the students. The investigation in the activity therefore had less of a, "do what I do" sense about it and required the students to think more for themselves about the process of manipulating variables.

However, evidence for autonomous learning experiences was not solely relied on in students' reports, but was also triangulated with observational data. One of the most clear cut observations of student autonomy was among a group of students in group Navy who completed the Recreate a Crater activity. Several students realised they didn't really know how the different options for composition of the impactor differed from one another in terms of their density (they are just labelled by name on the Impact Calculator by ice, porous rock, dense rock and iron). Instead of relying on the teacher to provide them with an answer, given that they had all the tools they needed to discover the answer for themselves, they decided to carry out an additional investigation alongside the activity to explore what happened when the only thing they changed was the composition of the impactor. Not only did the students record their data, they also drew up a graph to plot their results. These students displayed self-directed learning in that they worked independently of the main class and did not require any help

to do this (Birch and Ladd, 1997), their behaviour also coincides with what Lee and Anderson (1993, p. 594) describe as *intrinsically motivated to learn science*, they claim this is demonstrated by being naturally inquisitive about natural phenomena and actively attempt to make connections in their learning. These students not only did additional work to what was asked of them, they conducted an investigation that they came up with themselves, in order to further their understanding of the subject matter, therefore demonstrating a prime example of self-directed autonomous learning and successfully constructed their own understanding as explicated by constructivism.

Another group of students also in group Navy also did a similar thing. They had noticed that when changing the various parameters of the impactor, it always produced a circular crater. They reported to their teacher that they expected if the impactor was heading towards Earth's surface at a shallower angle, as opposed to directly vertical, it would produce a more elongated crater, bearing more resemblance to a tear drop as it grazed across the surface. With their hypothesis already formed, they continued to carry out an investigation. The boys discovered that their results were correct and explained that they had learnt that craters are always circular because the shape of a crater is caused by the explosive energy that is released in all directions when the object strikes the Earth's surface.

These two observations demonstrated younger secondary students (aged 13-14) not only carry out self-constructed, open inquiry investigations, but did not even need prompting from the teacher to do so. They had simply been carrying out the classroom activity, noticed something that triggered their curiosity and since they already had the tools they needed (the simulator) and the support from their teacher (i.e. he did not object to them straying from the lesson task), were able to expand their investigation and expand their knowledge through their own experiences of discovery.

However, despite these promising indications of autonomous experiences, there were some conflictive reports from students in the post-PPQ. One question asked students 'what did you like most about the activity?' and one common theme that emerged was named accordingly in response as 'Freedom and Student Control'. Ten students who participated in the D2E case study gave responses that correlated with this theme:

"Freedom"	(Male, group Pink)
"Making our own crater"	(Male, group Green)
"We got to create [an impactor] ourselves"	(Female, group Blue)
"We got to change/use asteroids by ourself and not be shown"	(Male, group Blue)
"I liked that we were all doing different things"	(Female, group Blue)
"That we could create our asteroid"	(Female, group Brown)
"The freedom"	(Male, group Brown)
"We didn't have too much help – only when we needed it"	(Male, group Brown)

"We got to find things out independently" "We got to choose what place we blew up" (Male, group Brown) (Female, group Brown)

Each of these comments reflects on a perception of working independently, making their own choices and having freedom in their learning. All of these are therefore considered to reflect autonomy.

However, an interesting contradiction is seen in the results to the question that followed, "what did you like *least* about the activity? A theme that emerged from students' responses here was titled "Limitations or Restrictions", seemingly the antonym of the previous theme.

This theme received nine relevant responses with regard to the Down2Earth activities, these were as follows:

"We could have done more investigative work"	(Female, group Green)
"We could investigate more"	(Female, group Green)
"We have restrictions"	(Male, group Green)
"Only being able to change certain things"	(Male, group Blue)
"Not needing to change anything else"	(Female, group Blue)
"Being restricted on what we can and can't change"	(Male, group Blue)
"I wish there was more choices"	(Male, group Brown)
"We only did one group"	(Female, group Brown)
"That we had to use certain variables"	(Male, group Brown)

Although this theme appeared to be the antonym of the previous "Freedom and Student Control" theme, when explored further, it became clear that these responses were in fact referring to separate aspects of the activity and did not necessarily present conflictive opinions among students, and instead reflected one another.

For example, three students gave responses that appeared across both themes. That is, their response for what they liked most about the activity applied to "Freedom and Student Control" and their response for what they liked least, corresponded to "Restrictions and Limitations".

A student from case Brown responded with "we got to find things out independently" to what they liked most and with "that we had to use certain variables" to what they liked least. Another student from case Brown gave these responses, "that we could create our own asteroid" and "we only did one group". Finally, a student from case Blue gave the responses "we got to change/use the asteroids by ourself and not be shown" and "only being able to change certain things". All of these students completed the Making a Meteorite activity.

When reviewing these responses together, it is clear that they do not correspond to the same aspect of the activity and in fact, given their responses to what they liked most, it is not surprising that a limitation in another area was what they liked least. Upon reflection it is considered that these students demonstrate the importance of autonomy and influence it can have over students' learning experiences. Where they have control over their learning, it stands out as the most positive aspect of a lesson, however where they are restricted in what they can do, it is notably the most negative element.

These responses are slightly problematic in that an experiment must be restrictive in some way. In order to investigate a parameter, you must control all other potential influences. The researcher considered that the best way to address the problem of restrictions in the activity would be to allow students to construct a research question for themselves and design the investigation process, i.e. carry out full or open inquiry.

This idea was echoed in the feedback from Mr Brown in his interview.

Researcher: If you were to run the activity again with those students, how would you run it differently?

Mr Brown: ... I would probably... separate them into... groups but I would also, in the questioning lead them up to what things are going to affect the size of a crater. Um... because with questioning you can lead them to the four things they were investigating and then if you give them a bit more ownership for it, then they'd be more likely to follow the instructions, as opposed to just giving them the instructions.

This response indicated that although the instructions were useful, it would be better to have them as a guide or supplement rather than a fixed prescription; students would consequently experience more autonomy in their learning if they were given this ownership. Mr Brown's students, Isabel and Jack also appeared to be of the same opinion:

Jack:	I think it helps us learn better when we have more control over, like over the practical	
	or over the work we're doing.	
Isabel:	I think instead of just like following instructions you've got you have to make your	
	own decisions and you don't want to make bad ones.	

It is revealed in multiple sources of data that although students were provided with an opportunity to work autonomously during the D2E activity, there is still room for improvements. On reflection on

students' feedback regarding the restrictions they perceived, it is considered that improvements to autonomous learning ties closely with improvements for differentiation. As well as redeveloping the Down2Earth OAs so there are three levels of differentiation, the lesson plan will also be altered to add in an initial discussion at the beginning of the lesson. This is in response to the comments from Mr Brown and is also considered to align with IBL. This initial discussion can also help the teacher to get a sense of students' current understanding of the subject area and will help them to better identify which level of the differentiated activity to implement.

10.4 Collaborative/Cooperative Learning

Group Blue was a particularly interesting group in terms of this dimension. If the initial observation of a day-to-day lesson with group Blue was anything to go by, it indicated a lack of group work in physics lessons. The lesson was very much didactic-based, with the teacher writing things on the board for students to copy into their books, and work was carried out independently. Students appeared to be highly disengaged, with high noise levels, little of which appeared to be lesson-related, and work output was extremely low.

However, during implementation of the Making a Meteorite activity, students were put into groups according to the variable they would be investigating and were essentially left to complete the activity, with the teacher removed from the focus of the lesson. On this occasion, conversations were significantly more task focused, the teacher did not have to prompt students as much (thus was able to take a more facilitative approach and students were able to work more autonomously) and they displayed a much higher level of engagement. Johnson and Johnson (1999) emphasise that cooperative learning has led students to learn more lesson material, report higher levels of confidence and motivation to learn and present more positive attitudes towards the subject. These effects were heavily evident in group Blue from the almost bipolar behaviour between occasions.

One group of male students in group Blue were in deep discussions relating to the task and about what they predicted might happen when they changed various parameters of the impactor. Similarly, another group of all male students were seen to be discussing what parameters they had already applied on the simulator and what had occurred in the outcome, others would then pitch in and explained how their outcome had differed when they had applied a different value for that particular parameter. This behaviour is reflective of idea discussion, debating and reviewing one another's work, all which have been shown to enhance cognitive engagement (Guthrie and Wigfield, 2000). Such discussions are also characteristic of IBL and are reflective of Dewey and Vygotsky's conceptualisations of learning as a social process.

Another noticeable characteristic of these students was that on most occasions, students would ask their peer *what* they had done or *how* they had done it, rather than plainly asking, what answer did you get? This behaviour was suggestive of cognitive engagement (Fredricks *et al.*, 2004) and prompted cooperative learning between peers. The students did not seek to merely complete the lesson task, but sought to understand the process and grasp the subject matter and again held relationships with their peers that were strong and positive enough for them to feel comfortable seeking help and guidance. This is considered to be reflective of a sense of relatedness among students which is considered to be integral to student engagement (Gagne and Deci, 2005). What was also encouraging from the observation of group Blue was that these occurrences of cooperative learning appeared to develop naturally among students, without any prompts from the teacher or researcher.

Nonetheless, despite the abundant literature praising the merits of collaborative and cooperative learning in the classroom, there are several requirements that must be met (see Johnson and Johnson, 1999). If not conducted effectively, attempts at group work can produce negative effects. In this study, it was noticed that there were occasions where group work provided an opportunity for students to slip into a more 'task avoidance' mind-set (Lee and Anderson, 1993), relying on their peers to do the work. For example some of the female students in group Blue, were seen to be directly copying the work from their peers instead of following the investigation process themselves.

The researcher therefore used a focus group with students from group Brown to ask them about their perceptions of group work in lessons.

Researcher:	How do you feel you work best in science? Maybe individually, or in pairs, or in groups?
Several Students:	Pairs or groups
Olivia:	It depends
Researcher:	What does it depend on?
Olivia:	It depends on what we're doing because sometimes people don't get it [the lesson task], then they want to work with other people but then if they do get it, they'll want to work solely.
Researcher:	Okay, that's great. And those of you who want to work in pairs or groups, is that because it's more fun to work in groups or do you find it actually helps you get the work done?

Alfie:	It's easier to work off each other, like some people might have stronger points in that subject so they can teach other people stuff
Isabel:	Yeah, they can kind of guide you
Mike:	It's easier to interact with others

Here the researcher not only wanted to understand students' perceptions of group work but whether these perceptions were rooted in a purely social perspective or whether they felt it benefited them academically. It was reassuring that the students commented on how they can work cooperatively to help each other out and Olivia even said that group work is not always preferable. She also drew on students' ability and if people are struggling, then they would want to receive help from their peers. This is reflective of the comments made by Johnson and Johnson (1999) whereby cooperative learning is seen to improve productivity and morale among students compared to working alone. Panitz (2002) also emphasise a different perception cooperative learning can offer when it comes to making mistakes. Where students work entirely independently, if one student makes a mistake it can feel like a limitation to that student's ability, thus dampening their confidence and self-esteem. However, when a mistake is made as a group, it acts instead as a learning tool, encouraging group members to work together to identify the solution to their problem or error. The researcher also recognises this final point to tie back to goal orientations. Where cooperative learning encourages mistakes to be taken positively and promotes effort from students this is demonstrative of a learning orientation, where students perceive failure as a learning tool and in response they promote their effort. This is alternative to performance learning where failure is assumed to be due to a lack of ability, and consequently expected to be more common in individual learning.

10.5 Context Novelty

The Down2Earth case study implemented a set of OAs that followed areas of the compulsory physics curriculum in key stage 3, but applied a context of extra-terrestrial impacts with Earth. It was clear from classroom observations that many student in the study had never experienced an activity like this in their classroom and it was therefore hoped that this novelty would successfully trigger students' situational awareness as suggested by Hidi (1990). The novelty element was also considered to be capable of bringing a 'wow factor' to the lesson, whereby students' emotional engagement dramatically peaked. Though this is something that can be easily demonstrated in chemistry, with trademark explosions and the mixing of chemicals, this is perhaps much less common in physics lessons.

The first line of evidence for students' interest in the D2E context was provided among their responses to the post-PPQ question that asked them what they liked most about the activity. Of the 109 students that answered this question, 48 of them (44%) gave a response that related to the context. This included responses regarding the Impact Calculator itself (16 responses), craters (15 responses), asteroids (6 responses), space/astronomy (3 responses) and blowing things up (8 responses). All 47 responses are provided in Appendix 15.

These promising results were successfully triangulated with feedback from a focus group with students from group Navy.

Researcher:	What did you think of the topic of space and impacts?
Tyler:	Interesting
Charlie:	That was cool
Matt:	You get to shoot rocks and stuff
Luke:	Yeah, it's a lot more interesting than likepower

Researcher:	How would you describe the activity that used the online simulator?
Matt:	Fun because I blew up my house
Tyler:	It was interesting to see like, how much it [an impact] can actually affect places

Though it is heartening to see that the topic of the OAs was well received, Tyler was also encouraging in his response that also reflected that he was interested in learning about the consequences of meteors and impacts, rather than reflecting purely on the element of being able to blow things up.

However the most authentic responses to the context which were also entirely unprompted were provided in classroom observations. Observations of group Brown revealed that although the lesson observed pre implementation was a practical, students' affective processing was noticeably different in the initiative lesson and the OAs appeared to present students with more of a 'wow' factor than in the day-to-day lesson. It was evident that students' favourable responses were largely directed towards the Impact Calculator simulation. While students were starting up their laptops, the teacher demonstrated the simulator to the students, which yielded an "*oh my god*!" and "*woah, that's cool*" from various students. A female student was also particularly animated in her response and when exploring the simulator and the Google Maps interface exclaimed, "*oh my god, it even has cars! You*

can see the cars!" once she had superimposed her crater on the surface she then exclaimed "*that crater is huge! I can't believe how big it is!*". Another female student also said out loud, though seemingly to no one in particular, "*this is amazing, actually*". It was promising to see that this excitement and amazement was shared between both male and female students.

It was also considered that where students identify a point of interest with something, their intrinsic motivation is increased (Lee and Anderson, 1993). Lucy in group Blue who has been discussed prior, was one of the first students to complete the D2E task, for comparison, in the previously observed day-to-day lesson, Lucy did not complete the set task. Lucy herself responded to this notable change in performance, turning to her peer to say said "*I never do any work in this class, look at my book there's pretty much nothing in it*". The researcher noticed as she flicked through the pages of her workbook, that there were only a few sentences on each page and in some instances just a date and title. Another girl nearby heard her make this comment and also followed up with, "oh my god I know, *me too. I've actually completed the work, I never do anything! Look at my last lesson, I wrote one sentence*", again also flicking through her workbook and revealing a similar amount of work. Then she followed this by turning away from Lucy and turning to the researcher and said, "*I actually did the work today*" and held up her book to show two full A4 pages of notes she'd produced in the D2E activity lesson. The behaviour of these girls' demonstrated increased attention to the lesson work, greater concentration, enjoyable feelings and increased effort; all of which are indicative of a situational interest (Krapp *et al.*, 1992).

Though this observation does not demonstrate that the lesson context was solely responsible for this dramatic change in the students' engagement, the format of the previously observed lesson suggested that the D2E activity was a novel experience and did not conform to their usual physics lesson experiences. Given that these students had performed inadequately in the previously observed lesson and faced no consequences, it is subsequently considered that they did not hold any extrinsic motivation to complete the lesson work, therefore it is assumed that these girls had taken a genuine interest in the initiative lesson and were again reflective of Lee and Anderson's case of "*intrinsically motivated to learn science*" (1993).

However, there was evidence indicative that some students did not buy into the activity. One group of students in group Navy appeared to be largely disengaged and unfocused on the task. Though it was disheartening to see that the activity had failed to capture the interest of these students, it was evident from the teacher's response that this disengagement was not an unfamiliar occurrence. One indication of this was provided when the teacher announced that if they did not finish their worksheet in the lesson, they would have to finish it for homework. This prompted the following dialogue between the teacher and one of said students:

Sebastian: Oh but we've already have loads of homework

Mr Navy: Well if you have a lot of homework what does that suggest?

The teacher thus implying that this student often failed to complete lesson work and subsequently was left with a large amount of homework. The initial trigger point for this student and several of his peers' disengagement is consequently unclear. One possibility is suggested by Stipek (2002) who emphasises that where teachers show confidence in their students' abilities then students in turn respond positively in order to prove themselves worthy of this opinion. In relation to the students in group Navy, it could be that due to their typical behaviour, teachers portray low expectations of these disruptive students and implement minimal consequence, therefore leaving the students with no immediate motivation to prove them otherwise.

This final point is indicative of the importance of the teacher's role, regardless of the classroom activity. Even during lessons that are intended to be primarily student-centred, like the D2E activity, the teacher still had a vital role and should maintain and reinforce their informal relationship with students. This leads on to the final dimension of the evaluation framework, Teacher Role.

10.6 Teacher Role

Much of students' learning experiences were considered to be largely influenced by the teacher. Even where lessons are primarily student-centred, it still remains the teacher who must allow for this and ultimately governs the delivery of curriculum content of day-to-day lessons. Therefore much of the points discussed, student autonomy, collaborative learning and differentiation, cannot occur without input and encouragement, or at the very least, permission from the teacher. Therefore, even in classrooms that are largely student-centred and involve high levels of student autonomy, the teacher still plays a vital role. Due to their influence over multiple dimensions and classroom aspects, the role of the teacher is the final dimension of the evaluation framework.

Teachers are integral to a classroom lesson, they are responsible for every student in their class and are the expert in the room. Teachers can have a large influence on students' attitudes, values, motivation and aspirations, both positively (Aschbacher and Roth, 2009; Ametller and Ryder 2014; Hamlyn *et al.*, 2017; Mujtaba and Reiss, 2013) and negatively (Bevins *et al.*, 2005; Gill and Bell, 2013; Hamlyn *et al.*, 2017).

Though the OAs were provided to teachers in the same format, it was evident that each teacher implemented them in their own way which was always slightly different to others. That is, no two teachers implemented the activity in identical styles or formats.

Five groups participated in the Down2Earth case study. Implementation in groups Navy, Blue and Brown was carried out by their teacher, and for groups Green and Pink, implementation was carried

out by the researcher. The implementation structure and the role of the teacher in implementation are described for groups Navy, Blue and Brown.

According to Grasha (1994), the teacher in group Navy appeared to employ a facilitative role. He possessed a very calm and relaxed attitude and his demeanour suggested that he was confident in his students' behaviour and ability. Mr Navy gave a brief introduction to and demonstration of the Impact Calculator and while students were logging on to the computers he initiated a classroom discussion around ideas of how to prevent an impact on Earth. However, as soon as students had been set the task and provided with their worksheets, Mr Navy took a step back. Although he was seen to be wandering around the classroom and interacting with students, he generally left the students to work autonomously. He was not imposing or checking over their shoulders, but instead was seen to be merely partaking in informal conversations with students and only assisting when they asked him. Grasha (1994) describes a facilitative teacher as one who emphasises a personal nature within their interactions with students and focuses on their needs.

The only other group in the D2E case study where the teacher showed a similar style was in group Brown who completed the Making a Meteorite activity. After introducing students to the activity, Mr Brown had no central role in the lesson again until the final 10 minutes where he organised a followup classroom discussion. He asked students from each group (depending on what parameter they had investigated) to report their findings and then asked students to think about what properties contribute to kinetic energy.

In group Brown, the researcher was able to observe two regular lessons of the co-ed group of students, prior to the initiative lesson, one physics and one biology lesson both taught by the same teacher who was in fact a Chemistry specialist. It was clear from these lessons that the teacher had a good relationship with the students and encouraged all students to have a voice. The lessons would typically begin with a class discussion about the topic they would be covering that day, this was perceived by the researcher as a way for the teacher to gauge an idea of what the students already knew and also to get them thinking about the scientific concepts they would be covering in the lesson ahead. The teacher would ask the class a question which would see a lot of hands go up, though often students would not provide the answer the teacher was looking for. However, on such circumstances, the teacher would remain encouraging with his response, instead of saying directly "no that is not right", he would reassure students with "you're on the right lines" or "that's a good idea but not quite what I was looking for". These same students would be seen to again raise their hands to answer questions, demonstrating that they had not been disheartened by not answering a question correctly and their confidence was not knocked. This process of guiding students through questions and encouraging them to explore various options is again indicative of a facilitative role (Grasha, 1994). The researcher also considers that these observations are suggestive of sense of relatedness and

competence among the students, two of the three basic needs set out by SDT which are essential in promoting engagement in the classroom (Deci and Ryan, 2012). Furrer *et al.* (2014) explain that relatedness is experienced when a teacher shows acceptance or warmth to an individual as demonstrated by Mr Brown's encouraging responses. Students' feelings of competence are promoted when there is clear structure to the lesson and when the teacher provides guidance and confidence in their achievements.

It is therefore perhaps not surprising that for the General Interest scale in the pre-PPQ, students from groups Navy and Brown reported the most positive scores among all the groups in the D2E case study. Though none of the items asked directly about the teacher, it is of interest that these students who experienced similar teaching styles had the highest general interest in physics.

These results align with those of Hamlyn *et al.* (2017, p.29) who emphasise "*having a good teacher*" as one of the most encouraging factors of learning science. These scores are summarised in Table 20.

Table 20 – D2E Case Study Group Scores on Pre-PPQ General Interest Scale	
Group	Mean Score (SD)
Navy	3.90 (0.59)
Brown	3.67 (0.77)
Pink	3.58 (0.83)
Blue	3.45 (0.70)
Green	3.14 (0.89)

It is also important to reiterate that Mr Brown was not a physics specialist, yet his implementation of the initiative activity still demonstrated positive results. This argues against the points highlighted in Chapter 3.3 by Ward (2018) and the National Audit Office (2016) as teacher specialism as a barrier to reform. This particular group demonstrated that where the teacher had a set of OAs readily available, implementation of an inquiry-based physics lesson was successful.

Some students from group Navy also participated in a focus group with the researcher, here they commented on the teaching style of Mr Navy and its positive influence on their lessons.

Researcher:	Are you able to name three good things and three bad things about your
	normal physics lessons?
Tom:	Our teacher is boring
	[all laugh]
Tom:	No, for a good thing, the teacher

Several Participants:	Yeah [in agreement]
Alex:	Like, he teachers differently to other teachers
Ben:	Yeah
Fred:	It's not like, textbook. We learn in different ways so you do actually remember
Tom:	He's more interactive
Fred:	He is interesting. Opposed to like, English or geography

Mr Brown and Mr Navy were seen to be prime examples of the positive effect that a teacher's role can have on student experiences as reflected in other literature. However, the D2E case study also presented an example of the negative impact the role of a teacher can have on students' experiences. This was observed in group Blue, year 9, co-ed group of students.

Two particular students in group Blue, Lucy and Jess, were seen to be amongst the most uninterested and disengaged students in the day-to-day lesson (see Chapter 10.1; 10.5). However, in the implementation lesson, they were the first to complete the set task. This was a huge achievement for them not only in terms of their performance but also in their classroom behaviour. However, instead of encouragement and praise from the teacher, Mr Blue responded with the comment, "*there is a state of shock and disbelief, you've actually done the work*" and "*are you feeling okay*?". Despite the lighthearted, joke-like tone of the comment, this could be perceived to undermine the students, suggesting that he had never expected them to complete the activity. This behaviour counteracts the points made by Furrer *et al.* (2014) who assert that student-teacher relationships and feelings of relatedness and competence are strengthened when the teacher expresses to students that they believe they have the ability to do well.

However, the teacher partly recovered from this initial comment when he followed it up with, "do you know what, I am hugely impressed with you guys, I've got to say" and "does this mean you might change your mind and take triple science next year?". Unlike the previous, these comments were much more encouraging to the students, here the teacher is expressing more of an indication that he does believe they are capable of the work and should consider taking triple science in GCSE, but it is the students who are unsure in the matter. Though, this uncertainty in the students is not surprising if they are often subjected to comments like the former.

However, where groups Blue and Brown were observed before implementation, this revealed a distinct change in teaching style of both teachers. In the day-to-day lesson, Mr Blue was seen to

display behaviour typical of traditional teaching methods, for example, writing content on the board for students to copy into their workbooks. In both lessons observed pre-implementation, Mr Brown ran practical activities. Characteristically, Mr Brown would gather students around one table where he would demonstrate the experiment and explain what they would be looking for and measuring. The students would then break off into groups to repeat the experiment.

Turning to the implementation lesson, Mr Blue although still provided students with a table template to record their results in, he almost entirely left the students to their own devices after the first ten minutes of the lesson. This meant that students had to work through the activity themselves and could not rely on Mr Blue to provide any information or knowledge.

Mr Brown, though having conducted practical activities in both lessons previous, appeared to take a greater step back during implementation of the OA. Here his demonstration merely consisted of showing students the Impact Calculator and how to alter the parameters. He did not at any point explain how to carry out their investigation or what to record.

Although observations of all groups within this case study demonstrated that teachers will employ an OA and deliver it in their own personal style with slight tweaks, such observations of Mr Brown and Mr Blue were implicative that despite such variations, what was seen consistently across groups was the lack of didactic approaches. That is, in all instances, implementation displayed fewer traits of traditional teaching styles and more traits in line with constructivism and inquiry. Although the role of the teacher is highly influential to implementation, the OAs were structured enough to prevent a didactic approach to delivery.

10.7 Overall Effectiveness

Having reviewed implementation of the D2E activity through the lens of the evaluation framework, an illustration to summarise the OA's overall effectiveness across individual groups is summarised in Figure 12. This continuum technique is adopted from Reeves (1994). When defining the individual dimensions of the evaluation framework in Chapter 5, each dimension was described as a continuum with two opposing conditions at each end. The points on each continuum therefore represent the position that is considered to be the most reflective of the evidential data from implementation for each group. This is to provide a visual summary of the mixed method data synthesised through the evaluation framework dimensions.

Each D2E participant group is illustrated on the continuum by the relevant colour and where there is overlap of positioning, points are vertically lined. These positions are by no means an upper or lower limit of the OA's potential and merely illustrates how implementation was observed under each group environment.


Figure 12 – Illustration of the Overall Effectiveness of the Down2Earth OAs

Chapter 11

Case Study 2: Hubble's Expansion of the Universe

The Hubble's Expansion of the Universe OAs were designed for key stage 5 students. The two OAs are not therefore not particularly comparable of one another and varied in terms of their curriculum applicability, target age group, content and theme. Though the Hubble activity was again differentiated and the variety of implementation methods were described to teachers, as with the D2E activities, it was at the discretion of the teachers how or if they use the material provided.

11.1 Goal Orientation

Students' orientations in this case whether learning or performance focused were less observably pronounced than in the D2E case study. Unlike the Down2Earth OAs, the Hubble activity had an accompanying introductory PowerPoint presentation and lesson plan for the teachers. The presentation was intended as a tool to 'set the scene' for the students, introducing them to the subject content and how the data they will be working with is collected and interpreted. This takes students back in time to Hubble's discoveries and his first plot of supernova targets.

The introductory PowerPoint also set out the main objectives for the Hubble activity:

"You are going to be looking at some real supernova data in order to produce your own Hubble plots"

"From your plot, you will be able to calculate the rate at which the Universe is expanding and how old it is"

However, as the use of the material provided in the Hubble activity was at the discretion of the teacher it was possible that they would not make use of the PowerPoint. This was the case for group Red and as a result, although the teacher briefly introduced students to the lesson activity, no clear objectives were set out at the start of the lesson.

In response, the Hubble activity will be adjusted so that learning objectives are also set out at the beginning of the Student Worksheet and in the teacher's Lesson Plan. The latter is particularly

important as even if the teacher doesn't provide any handouts to their students, the teacher will immediately see the activity objectives and can convey this to their students. It is acknowledged that the learning objectives also should have been broken down further and reinforced throughout the activity. This is in alignment with Quinn's recommendation that content should be presented clearly and frequently (1996).

Focusing on objectives at the beginning of a lesson presents an opportunity to induce discussion and prompt students to think about the processes behind the given concepts and equations they encounter in their physics syllabus. Where typically these students will follow a specific syllabus they must learn in order to pass exams, experiences of the researcher were indicative that often there is little consideration to how such understanding originally came to be and how science evolved into the knowledge held today. The introductory PowerPoint in the Hubble activity goes through this history and introduces students to the evolution of astronomers understanding of our Universe.

By starting the lesson with asking students to come up with some ideas around how they might attempt to calculate how quickly the Universe is expanding or how they could determine its age, it forces students to think about the background science around the subject matter and how such phenomena might be investigated.

The researcher therefore considers that not only do learning objectives help students to focus their attention (Quinn, 1996), but they can also prompt classroom discussions and trigger students to ask questions and inquire about how our understanding of particular scientific concepts came to be. The influence of this novelty that extends beyond the prescriptive syllabus, on students' learning experiences is discussed in Chapter 11.5.

However, although the researcher identified several ways the lesson objectives could have been clearer and also used to induce discussion, there is evidence to suggest that most students felt confident in their work and what was expected of them. Chapter 9 highlighted how students' reported significantly higher confidence in relation to the activity than towards their day-to-day physics lessons. Mrs Yellow reported in her interview that she believed this could be due to the clear goals of the Hubble activity.

Researcher: So you may remember I gave out questionnaires last year before they did the activity and after, and I have actually found that their confidence they report towards the activity is higher than what they report for their day-to-day physics lessons. Do you have any idea what might be the reason for this?

Mrs Yellow: Um, probably because in the activity it's very clear what they're trying to achieve and in their physics lesson you're not only trying to achieve the objectives in the lesson you're trying to meet the syllabus at the end and there's nothing more depressing doing a lesson on something and then doing an exam question and not being able to relate the exams to the lessons, because quite often you can't.

The researcher also wanted to clarify that the reason for students heightened confidence wasn't simply because the activity was too simple and not challenging to the students.

Researcher: Okay perfect, so it wasn't that perhaps the activity was too easy for them and instead it was just very clear?

Mrs Yellow: No I don't think it is too easy for the kids, this is one that gives them perspective on it. It's not complicated, you know, I know you gave me the offer of doing something with light curves and photometry and I actually wish now I'd probably taken you up on that because I probably would have extended it a bit, but it really does depend on the amount of time I've got.

This final comment from Mrs Yellow is referring to an extension of the Hubble activity where students go a step further and obtain the supernova photometric data for their light curves. Mrs Yellow clearly values the importance of these extra activities that demonstrate to students where the science comes from and how this can give them a "*perspective*" of the science they learn in school.

The element of whether the task could have been too easy for students also leads us on to the following framework dimension, Differentiation.

11.2 Differentiation

When designing the Hubble activity, the researcher made a conscious attempt to provide several suggestions for implementation based on varied levels of guidance provided to students. This was done in order to account for individual differences among students and also the time available for the activity.

Teachers were seen to implement different levels of the activity to their students, but again, no teacher (or the researcher) provided within class differentiation. However, there were two particular groups within the Hubble case that perhaps would have benefited from differentiation. This is groups Black and Purple. Group Black could have perhaps benefited from within group differentiation as participants included year 10, 12 and 13 students, therefore at very different levels in their education and with presumably varying physics knowledge. Group Purple on the other hand, although all students were year 12 and had chosen science A-levels, they had not all chosen to take physics, so again, some students' participation in physics had ceased at the end of GCSE.

Feedback from students in the PPQ is also demonstrative that some students faced difficulties in the activity. PPQ responses to what students liked most and least about the activity revealed that some students did experience some difficulties. The relevant responses are summarised in Table 21.

Table 21 - Hubble Case Study Students Tost-11 Q Responses multative of Experiences of Differing					
What did you like <i>most</i> about the activity?	Group	What did you like <i>least</i> about the activity?	Group		
It was challenging	Purple	The graph because the scale was difficult	Purple		
How simple it was to understand	Purple	I found it difficult to plot the data on the graph	Purple		
Challenging – old school calculations	Purple	Difficult to create the graph on paper	Purple		
		Difficult to plot the graph	Purple		
		Initially it was hard to understand what to do, however this was quickly rectified	Red		
		Complicated numbers	Yellow		
		Using large varieties of numbers which proved difficult to plot	Purple		

Table 21 – Hubble Case Study Students' Post-PPQ Responses indicative of Experiences of Difficulty

Here it is evident that most responses came from students in group Purple. This is perhaps not surprising given that not all students were studying physics, however several other additional reasons are also identified:

1/. Implementation for group Purple was carried out by the researcher who was entirely unfamiliar with these students and was unaware of their learning preferences and abilities

2/. This was the largest group in the case study with 60 students, double the normal class size and therefore had the lowest teacher/instructor to student ratio across groups.

3/. This was not a standard class of students and was a group of gifted and talented students invited from various schools across South East Wales.

4/. There was no computer access and so consequently more restrictions on implementation methods

However, among those who did experience difficulties, the general consensus appeared to apply to the graph construction component of the activity. Given that the values for radial velocity ranged from 2,091 km s⁻¹ to 94,720 km s⁻¹ but with many points positioned towards the lower end (see Figure 13), it presented difficulties in scaling the graph when working on paper. The teacher in group Yellow however, identified that this would be problematic for her students and so provided them with a scale to use.



Figure 13 – Example of the Resulting Hubble Plot Produced in the Hubble Activity

Furthermore, there are no comments about difficulty from students in group Black. It is expected that this is because they were a particularly small group, the lesson was delivered by the researcher and each student had access to a computer. Both the researcher and the teacher were also on hand to support if needed.

However, what is also evident from students' responses is the contradiction in experiences of the difficulty of the task. The level of the activity appears to have been most suited to the students who reported that what they liked most was that the activity challenged them. Although another student reported the simplicity of the activity as what they liked most, there is a risk that tasks too simple will disengage students (Ofsted, 2013). Also, students who reported that they didn't like how difficult parts of the task were, were also at risk of disengagement (Fitzgerald *et al.* 2015a) and knocks to their confidence. Rittmayer and Beier (2009) highlight that where students fear failure, they are more likely to doubt their abilities. These responses are reflective of the importance of the recommendation made by Rittmayer and Beier (2009) who profess that lesson activities should be challenging but not impossible.

It is also interesting to note that although initially it was expected that differentiation would be problematic for group Black, no students in this group reported any difficulties. On reflection it is considered that this could be due to the small size of the group. With just eight students and both the teacher and the researcher present, support was readily available to students whenever they needed it. This was much more difficult with the 60 students in group Purple. In order to probe this dimension further, we turn to students' responses regarding their confidence during the activity as these are indicative of how difficult or simple students found the task. As with the D2E case study, we assume an optimum score to be 4.00 (within the range of 1.00 to 5.00).

Students who participated in the Hubble case yielded a mean Confidence score on the post-PPQ of 3.70 ± 0.63 which was revealed in an independent sample t-test to be significantly higher than what was reported on the pre-PPQ in relation to their day-to-day physics lessons (3.37 ± 0.79) , t (203) = -3.321, p = 0.001. Students therefore felt higher levels of confidence when completing the Hubble activity and were closer to the optimum score of 4.00.

Figure 14 illustrates the distribution of students' Confidence scores on the pre- (left) and post-PPQ (right). Here it can be seen that no student scored less than 2.20 (where 2.00 indicates strong disagreement) on the combined scale on the post-PPQ. The same however cannot be said for students' confidence towards day-to-day physics lessons where highly negative perceptions were more frequent.





Though only speculative, the smaller range in student scores on the post-PPQ and the greater percentage of students scoring 3.00 or above (90% in the post-PPQ compared with 73% in the pre-PPQ), could be indicative that the activity was more suited to individual differences and learning preferences than many day-to-day physics lessons. Had all students' yielded scores between 4.00 and 5.00 then it would suggest that the activity was much too simple for all students, however given that

there is a range in scores that are skewed more towards the positive end in the post-PPQ, it is suggestive that many students felt more capable in completing the task.

Qualitative evidence of the positive impact that challenging students at the right level can have were revealed in observations of particular students. One example includes two female students in group Purple. These students were not studying physics A-level and therefore would not be expected to learn the content of the activity or be as familiar with the relevant physical concepts. It was apparent from their behaviour that the students saw this as a disadvantage and often sought reassurance from the researcher they were doing things correctly. However, this assurance was not necessary as they gained a good grasp on the activity. The two were seemingly unhappy with their values for distance as they felt that there shouldn't have been such great differences across the supernova targets. At the end of the implementation, the researcher retrieved her answer sheet and went over to the students to prove to the students that they had done their calculations correctly. One of the two students, Lauren, responded with "ah, I knew we must have", indicating a confidence regardless of the difficulties they faced and perhaps her fellow peer not being quite as optimistic. As she continued to scan the answer sheet and realise they had done it correctly, she looked at the researcher and said, "I should have done *physics*". This response was equally as heartening to the researcher as it was frustrating. Lauren was a prime example of a student who had the ability to perform well and pursue physics, but had chosen not to and whose behaviour indicated she had perhaps underestimated her abilities. If her compulsory physics lessons in earlier stages of education had been differentiated effectively to account for individual differences among students, or had she been encouraged by teachers to pursue physics, it could be that she would have chosen her A-level options differently.

Disbelief from students about what they had achieved was also observed in group Black. Several year 10 students participated in this group, again outside of the target audience of the Hubble activity. Though requiring more assurance and guidance than their older peers in the group, they were still able to complete the activity in the lesson time. Upon completing his calculations, one year 10 student, Harry, turned to the researcher and said, "*what, so that's the age of the Universe?*" signalling towards his final calculated value. The researcher confirmed this, to which he smiled and said, "*I feel like Einstein*".

These two experiences of students in group Purple and group Black are demonstrative of the impact just one lesson can have on a students' learning experiences. Both Harry and Lauren's comments are demonstrative of the confidence and competence (Gagne and Deci, 2005) they felt in response to completing the activity which has shown to be positively related to interest and engagement (Linnenbrink and Pintrich, 2003). The same authors also argue that confidence is much more malleable (given that it is task-specific) and subject to influence than individual interest which is more stable and withstanding. Given that the difficulty of physics is commonly reported by students as a deterrent (Williams *et al.*, 2003; Tripney *et al.*, 2010; Gill and Bell, 2013), it is considered that targeting students' confidence and confidence by effectively differentiating activities could therefore offer a strong starting point in promoting student interest and engagement in physics.

The researcher considers differentiation among students to be equally important at A-level as it is in other years, as students are not put into sets. Although many sixth forms and colleges require students to achieve a minimum grade in order to pursue a subject into A-level, students still present differences in their abilities and preferred learning styles. Though group Purple was not a standard class of students, they were all considered to be 'gifted and talented' with similar predicted grades across the sciences. However, the variety in the responses from these students is demonstrative of the importance and necessity of differentiation and how this was not accounted for efficiently by the researcher in this implementation.

Perhaps a better example of where differentiation is necessary within sixth form classes is group Orange. This was an A-level physics group of six students who displayed clear discrepancies in ability and confidence.

Almost immediately apparent in the observation of the implementation lesson was students James and Holly as the frontrunners of the class. They sat at the front, were the most vocal and were seen to confer with the teacher substantially more than the other students. Disheartening was the behaviour of the other students in group Orange who appeared to be exceedingly introvert in comparison to Holly and James. Very few observations of dialogue were recorded for these students throughout the whole two hour lesson and when they did speak, they were much more quiet and reserved. It was clear on several occasions that all three remaining students were confused by or struggling with the set task (participant 5 left the lesson before the main activity started). However, where Holly and James had been seen to check things with the teacher, as did most students across other groups in the Hubble case study, these three students appeared to remain quiet. At no point were these students seen to ask for help, but instead would wait until the teacher came over to them to check how they were getting on. After some 20 minutes into the activity, no students were seen to have provided Holly with any results and did not seem to be making much progress. In the case of one student, Tony, was seen to walk over to Holly and James' desks and merely look over their shoulders to see what they were doing – without saying anything. It was also noticed than neither Holly nor James reacted to this, suggesting that this could be a common thing that they are used to happening. This lack of acknowledgement of their peers also suggested that Holly and James themselves saw themselves as separate from the rest of their classmates. This action from Tony also revealed that fellow students also recognised the two students as more capable, not just the teacher and the researcher.

The necessity for differentiation in this group was overwhelmingly obvious. Where Holly and James flew through the activity and did not once seek any help from the teacher, the other three students

struggled tremendously and did not grasp the process of the task. What's more, James was seen to interrogate the astronomical database, Simbad, in order to obtain the radial velocity values even though the teacher had provided them with the table of values, thus demonstrating an additional ability and motivation in inquiring further.

However, teacher interview feedback in this case study was more positive regarding the appropriateness of the activity than that received in the D2E case. For this case, Mrs Yellow was interviewed. She had implemented the Hubble activity on consecutive years to her year 13 physics students, and thus had repeated implementation.

Researcher:	Do you think there were any students in particular that would not benefit from this activity?
Mrs Yellow:	In upper sixth?
Researcher:	Yes
Mrs Yellow:	I think if they're doing physics A-level, I think every student should do this.
Researcher:	Do you think the material was accessible for all the students?
Mrs Yellow:	Yes, oh yes.

Though this comment does not directly relate to differentiation, the teacher is still stating that she believes it is a useful activity for *all* A-level physics students. This question offered Mrs Yellow the opportunity to critique the activity and offer suggestions for improvements but that was not taken.

Overall, it is considered that the Hubble activity was more malleable and easier to differentiate than the D2E activities, perhaps because the different versions were more readily available than in the D2E OAs where differentiation was merely suggested to teachers. However, carrying forward the need for clearer objectives highlighted in the previous dimension, it is considered that this will also aid in the differentiation of the activity.

11.3 Student Autonomy

When designing the activity for the Hubble case study, it was immediately evident to the researcher that this OA would have to follow a more prescribed process than the D2E activities. However, though less autonomous than designing their own investigation, generating hypotheses and manipulating variables, the students were still able to conduct an investigation, interrogate and

interpret data, all elements that are in line with IBL. The researcher therefore still considers this as more autonomous than a didactic lesson where students are just presented with a series of facts they must learn.

PPQ responses revealed students' level of agreement with the incorporation of IBL in the activity. In the pre-PPQ students were asked to indicate their level of agreement with the statement 'we often carry out investigations in physics lessons', in the post-PPQ, this was rephrased to 'we carried out an investigation in the activity'. Students reported significantly higher levels of agreement with the post-PPQ item (3.88 ± 0.86) than with the pre-PPQ item (3.10 ± 1.01) , t (207) = -6.003, p < 0.001. Of the Hubble case study students, 77% either Agreed or Strongly Agreed with this statement and just 7% Disagreed or Strongly Disagreed (the remaining 16% responded with Neither).

Additional evidence is also received from students' responses in the post-PPQ to what they liked most about the activity. Responses from the Hubble case study revealed seven common themes, whereby a theme was considered present if three or more students gave a response related to the same aspect. One of these themes was subsequently titled 'Independence' and received responses from six students:

"Arriving at the age of the Universe using values I created"	(Male, group Orange)	
"The independence"	(Male, group Purple)	
"I liked finding out own results and maths"	(Female, group Purple)	
"The independent part – we weren't just talked through it"	(Female, group Purple)	
"The freedom of finding things out on our own"	(Male, group Black)	
[abbreviated] "It was fun to do the calculations ourselves. Also it	(Female, group Purple)	
was a sense of achievement to correctly calculate!"		

Each of these responses in some way describes an aspect that the students felt they did themselves and had control over and indicate that these students' need for autonomy set out in SDT were fulfilled, which in turn is seen to promote engagement (Gagne and Deci, 2005). Such responses were also considered to be relatively unprompted as the question was largely open ended and did not lead students to their answers in any way. It is promising that despite the structured process of the activity, and necessity to follow instructions, students still felt a level of independence in their learning.

Though only an assumption, it is possible that such responses were indicative of a new experience among students. If autonomous learning was something they often experienced in lessons it could be that they would be less likely to report this aspect as what they liked most about the activity as it would be a common occurrence. In particular the "*we weren't just talked through it*" indicates that this is the typical experience of physics lessons and is descriptive of a didactic approach.

As the activity involved students taking their own measurements, performing calculations and interpreting results with substantial room for error, generally, students within a group would not all obtain the same value for H_0 and their approximate age of the Universe. Though they did typically yield answers between 13 and 15 billion years which was considered an acceptable answer to arrive at. However, it is believed that such variations in results between students could have provided them with a a sense of ownership over their work. The particular process they followed in the investigation influenced their final results. This was evidently experienced by the student in group Orange who reported the first comment listed above.

Upon group Black's completion of the Hubble activity, one student said to the researcher, "*can we keep these*?", indicating to his worksheet, as soon as the researcher confirmed he could, he said "*ah, yes! I'm going to show this to my dad*". One of his peers also commented, "*that's the age of the Universe*... *if I showed this work to my mum she'd be like, what the hell*?". Both of these comments indicate that the students considered the work they had achieved in the lesson to be a big deal. They demonstrated a sense of pride and were eager to show their parents what they had achieved. It is believed that an individual would not feel a sense of pride without also a sense of ownership and without feeling like they themselves had worked for this achievement.

In order to further probe the opportunity offered by the Hubble activity for autonomous learning, a larger student voice is required for focus groups. Unfortunately, one of the greater limitations of the Hubble case study is that no focus groups were carried out. The primary reason for this was that the content of the Hubble activity was always the final topic taught in year 13, therefore upon finishing, students had completed the syllabus and were into revision before their final A-level exams. It was therefore not a convenient time to take students out of lessons at such a vital point in their education.

11.4 Collaborative/Cooperative Learning

Though the previous dimension has discussed the opportunities for autonomous learning provided by the Hubble activity, we now address the opportunities; or lack of, for cooperative and collaborative (C/C) working among students.

The occurrence of C/C learning in the Hubble activity was immediately evident from students' PPQ responses regarding what they liked most about the activity. Not only did this confirm students' perception of its occurrence but also their enjoyment of it.

Collaborative learning; or as it was so coined, 'Group Work' was another of the seven themes that emerged from students' responses to the post-PPQ question that asked students what they liked most about the activity. The responses were eightfold:

"Working together as a class"	(Female, group Yellow)
"The teamwork aspect of working together on an interesting topic"	(Unknown, group Purple)
"Working in groups"	(Male, group Purple)
"Being able to work together through the steps to work out the eventual answers"	(Male, group Purple)
"Working as a group on a new task"	(Male, group Purple)
"The maths aspect, group work"	(Male, group Purple)
"Interactive and interesting, finding out things in a group"	(Female, group Purple)
"Group work"	(Female, group Purple)

Though this data is evidence for students' positive perceptions of collaborative learning, and is supportive of Johnson and Johnson's (1990) point whereby it can enhance student satisfaction, there was also evidence for its benefit to students' learning. This was most apparent among students in group Black who although were all from the same school, they were from different year groups.

Student interaction in group Black was seen to be more cooperative rather than collaborative. Though working primarily individually (as had been decided by their teacher), where everyone was completing each step of the Hubble activity, often students were seen to turn to their peers to seek help or to offer it. This was most typically seen between Ellie (year 12) and Harry (year 10). Ellie was a very able student and flew through the Hubble activity. Harry had noticed her ability and trusting this, would turn to her if he was stuck which Ellie was happy to abide to. Though it might appear to some that this interaction was more beneficial to Harry than it was to Ellie, researchers have pointed out that the student of higher ability benefit from the process of explaining and demonstrating and the student of lower ability benefits by experiencing an approach modelled by their peer (Panitz, 2002). Kagan (1986) also highlights that where students help one another in lesson tasks, a supportive community builds which has been seen to raise performances of all members.

Harry was also very openly and expressively impressed with Ellie's ability, commenting on it on several occasions. Ellie was the first in the group to complete the Hubble activity, to which the researcher announced "*we've got our first accurate calculation of the age of the Universe*". The group responded with a cheer and Harry responded to Ellie saying she was "*the smart one*". These actions were seemingly building positive interpersonal relationships among students who would not typically interact with each other being in different year groups. Reflective of other research was the evidence of productivity, high morale and commitment to peers (Johnson and Johnson, 1999). It is also these kinds of interactions that work to build students' self-esteem, motivate them towards the learning process (Johnson and Johnson, 1990) and build a supportive community (Kagan, 1986). This process of social-persuasion where positive feedback and support is received from significant others (i.e

peers) is considered to be one of the primary foundations of building confidence (Bandura 1977; Rittmayer and Beier, 2009). The early educationalists, John Dewey and Lev Vygotsky conceptualised learning as a social process and that it is through collaborative efforts that learning occurs. Studies of peer acceptance among students have consistently demonstrated a correlation with high achievement (Roeser *et al.*, 1996; Anderman, 2002).

Where students in group Black demonstrated the benefits of cooperative learning, students in group Purple seemingly reaped the benefits of collaborative learning. The collaborative element in group Purple was entirely student-directed where no suggestions were made by the researcher. However, this was entirely unexpected given the large cohort of 60 students from various schools. Groups of students varied both in size and gender, some had between two and seven students and some were single-sex and some were co-ed. However, what was overwhelmingly consistent was the collaborative style and distinct roles of individual group members. In the majority of the groups, each student had a specific role in each stage of the activity, ensuring everyone was involved and did not take a back seat during the activity. In other Hubble case study groups where collaborative learning took place, typically each student would do each step of the process for a fraction of the data and then results were combined for the final calculations. However, in group Purple, where one student would be identifying a supernova's peak magnitude, another student would be calculating its distance modulus, another would be identifying its radial velocity and another would be plotting its point on the graph.

Whether these roles were assigned by one person within the group, decided democratically or just emerged organically was unclear and most likely varied from group to group. However what was evident was the all-encompassing involvement that this created, providing everyone with a distinct, important role that promoted a sense of belonging, encouraging motivation and commitment to tasks (Osterman, 2000).

Though these observations of group Purple were demonstrative of strong team work and efficient collaborative efforts among students, from a learning point of view it is considered likely that students would have sacrificed certain elements of the investigation process and thus, knowledge. By being in charge of a specific, singular role in a multi-phase process, it would be harder for these students to see the bigger picture and the overall process of the activity. They may gain mastery experience at their specific role but at the expense of other skills. Therefore their understanding of the subject matter as a whole would be incomplete. Where data was divided among students and combined later, this allowed each student to perform each step of the process and though less efficient, from a learning perspective it was considered more beneficial.

Having discussed the benefits of collaborative learning and highlighting that this technique must be implemented in such a way in order to be most beneficial to students, we turn now to the negative impact of neglecting the opportunity for collaboration. A particularly profound example of this was observed in group Orange.

Evident almost immediately was the notably more disjointed class of group Orange compared to other groups in the case study. The teacher had sought to implement a collaborative element to the activity by dividing the data among students to later be combined to produce a master Hubble plot. However, this did not go as planned and for a group task, there was almost no communication between the students. Instead students Holly and James worked collaboratively as a pair and the remaining students mostly struggled on their own. Communication only occurred between Holly and James and they were also the only students in the group to successfully complete the task.

This demonstrated the importance of effective communication and the efficiency of assigning each other distinct roles. Two participants from this case surrendered next to no verbal communication, and data collected on them in the observation phase was primarily non-verbal.

It was evident that there was a distinct range in abilities among students in this group and thus they would have benefited tremendously from effective cooperative and collaborative learning. Had Holly and James behaved as Ellie had done in group Black, helping her less able peers, and had the less able students behaved like Harry, communicated with their more able peers and learnt from their approaches, it likely would have been a more beneficial experience for everyone. What's more, not only were the other three students unable to complete the Hubble activity, but they also appeared to suffer socially. There was a distinct lack of a sense of community in this small group and given the deficiency of communication and spread out seating plan of the three students it was assumed that the students did not feel connected to one another, thus their sense of relatedness and one of their three basic needs in self-determination theory was not fulfilled (Gagne and Deci, 2005).

11.5 Context Novelty

The Hubble activity addressed specific astronomy content within the curriculum, therefore the context was not considered to particularly novel as it was something students would have encountered regardless. Instead, the activity was considered to possess a novel element in terms of the investigation aspect of the activity. Where students may be lucky enough to conduct investigations and practical lessons in some of their physics syllabus, for obvious reasons, the researcher considers that it is unlikely this opportunity would present itself for the cosmology and Universe content. Therefore, given the content of the activity, the expansion rate and age of our Universe, the researcher considers it fair to assume that many students; and even teachers without an astrophysics background, would assume this to be a topic that cannot be studied through practical lessons.

Though this was an initial assumption made by the researcher, it was also reflected to an extent by some of the teachers. Mrs Orange reported that "*all the astronomy content we've done so far has just been maths, but they've been struggling with past exam papers. So it was nice to get them doing something more applied and practical and I think it will help*". Similarly the unawareness of the opportunity of the activity was echoed by the teacher from group Black, "*this is great that you're actually able to use real data and use it to calculate the age of the Universe*". It was therefore anticipated that the opportunity for students to explore real data which they can use to calculate the expansion rate and age of the Universe in a classroom physics lesson, would be novel and unfamiliar to students. It was hoped that this divergence from student expectations could stand as a trigger for their engagement.

The first evidence for students' reactions to the activity was again provided in their responses to what they liked most about the Hubble activity in the post-PPQ. Another one of the seven themes among students' responses was so-called 'Working with Real Data'. This theme yielded responses from 12 students which were as follows:

"Understanding real life values"	(Male, group Orange)	
"Arriving at the answer of the age of the universe using values I	(Male, group Orange)	
created"		
"Being able to calculate the age of the universe from real data"	(Female, group Orange)	
"Hands on approach when handling data"	(Male, group Orange)	
"It was something new and different"	(Male, group Yellow)	
"Calculations and real life investigations"	(Male, group Grey)	
"The real life aspect"	(Male, group Red)	
"It was interactive, using real data. This makes it feel more relevant"	(Male, group Red)	
"Something different and interesting"	(Male, group Red)	
"Use real data to come up with an age of the Universe"	(Male, group Purple)	
"Using REAL data"	(Male, group Purple)	
"Being able to do something practical and real-life"	(Female, group Purple)	

Although each of these responses indicates the value students felt towards the opportunity of using real data, it does not necessarily indicate the novelty. Observations of student behaviour during implementation on the other hand, were able to reveal more about students' reactions. In group Purple, one student's response indicated that he had not realised that they were using real supernova data, when informed, his response was "wow, that's really cool. We usually just get made-up data, it's nice to be working with real science". This comment also demonstrated that the activity was a

deviation from normal physics lessons and what they 'usually get' and is also considered indicative of a 'wow factor' discussed in the D2E case study (see Chapter 10.5).

The process of calculating the age of the Universe involved several unit conversions, the last of which requires converting the age of the Universe from seconds, to years. Ellie in group Black however also appeared to find calculating the age in seconds novel, responding "*wow, that's in seconds?* … *Well I've learnt something cool*". Similarly the comment from Ellie's peer, Harry, cited earlier "*what, so that's the age of the Universe? I feel like Einstein*". Another male student in group Black also reacted at the end of the lesson with "*That's the age of the Universe.*.? *If I showed that to my mum she'd be like, what the hell?*!" Where all of these comments started with a question, they indicated a level of surprise and disbelief, the former of which has been seen to enhance student interest (Palmer, 2009; Hidi, 1990).

Another aspect of the Hubble activity that is perhaps not always evident in day-to-day physics lessons given its lack of presence in the curriculum, but despite having shown benefits of its inclusion (Gooday *et al.*, 2008), is that they were introduced to the background of the science. That is, rather than learning that the Universe expansion rate is x and its approximate age is y, students were in fact introduced to the process behind this knowledge and how science has evolved our understanding and how these values are calculated.

Learning *how* the expansion and age of the Universe is *calculated* is only really specified in the Edexcel curriculum which states:

Understand how astronomical distances can be determined using measurement of intensity received from standard candles (objects of known luminosity)

 $V = H_0 d$ for objects at cosmological distances

It is also suspected that many physics teachers, unless they have a background in astronomy will not make the connection between supernovae as standard candles and a Hubble plot. WJEC, OCR and AQA also merely request students learn Hubble's law and how to approximate the age of the Universe.

However, some students reported the procedural and historical components as what they liked most and also what they learnt from the activity. Below are some examples of student responses from the post-PPQ about what they learnt from the activity:

"how people worked out the age of the universe"

"how we know how old the universe is"

"using results to find a value for something which can be used to find another"

"I learnt more about the history of scientists trying to discover the age of the universe"

"how to do important calculations that give the age of the universe"

"I learnt about the universe and how numbers I take for granted are calculated"

"how investigations are carried out and how conclusions are formed"

"I learnt a scientific method using general mathematics to find out the age of the universe"

(In order: Male, group Orange; Female, group Purple; Female, group Purple; Female, group Purple; Male, group Purple; Male, group Purple; Male, group Black)

A final piece of evidence comes from a particular item in the PPQ, though it is not directly related to novelty, it asked students to reflect on the excitement they experience in their day-to-day physics lessons and in the activity. An independent sample t-test demonstrated that students perceived the Hubble activity to be significantly more exciting (3.73 ± 0.79) than their day-to-day physics lessons (3.30 ± 0.90) , t (208) = -3.674, p < 0.001.

11.6 Teacher Role

The Down2Earth case study emphasised the integral role of the teacher that remains even when implementing a student-centred activity, and the same also applied to the Hubble case. Many points made in previous dimensions have touched upon or indicated the importance of the teacher's role. We again discuss this dimension last as the success and effectiveness of many of the previous dimensions is largely influenced by the teacher.

Of the six groups who completed the Hubble activity, three were delivered by the teacher, groups Red, Orange and Yellow and three were delivered by the researcher, groups Purple, Grey and Black. Groups Purple and Grey were the twilight workshop sessions with students from different schools, however for group Black, though the researcher delivered the activity, the teacher remained present and interactive with the students.

In the eyes of the researcher, group Yellow by far demonstrated the most successful implementation by the teacher. This was also the only teacher (known to the researcher) who repeated implementation with her next year 13 students the following year. Observation of both occasions of implementation was incredibly interesting and valuable to the researcher. Though Mrs Yellow's role per se did not differ between occasion, her confidence in delivering the activity was noticeably higher and this, in effect, seemed to also benefit students' learning experiences. The teacher herself had picked up on the difference, at the end of the second implementation she reflected, "*I will definitely be using this again next year, because I'm getting better each year. I know this works and they are really engaged aren't*

they". This observation is reflective of findings from Fitzgerald *et al.* (2015a) who found greater improvement in students' content knowledge and perceptions of science at school subsequent to initiative on occasions where teachers had previously implemented the same material.

In feeling more confident in her understanding of the activity, Mrs Yellow also implemented a more advanced version on the second implementation, demanding more work from the students. The first year, students used the automated spreadsheet to calculate the distances to the supernova targets, however on the second year, the teacher had students apply the distance modulus equation themselves to calculate the distance. The teacher was also seen to make more connections between the activity and other areas of the curriculum they had covered. For example she asked the students the definition of radial velocity and how it is calculated. Students were able to connect this with Doppler shift.

Mrs Yellow had the students work collaboratively by dividing the supernova targets among them. Though this collaboration was not student-initiated as it had been in group Purple, Mrs Yellow's technique of collaboration enabled all students to experience each part of the investigation, allowing students to get the most out of the activity. It is considered that Mrs Yellow took on a highly facilitative role in her implementation, though students were responsible for working through the investigation, the teacher ensured they understood what they were doing at each stage and organised them to collaborate efficiently and also built in guidance and support where they needed it. Quite opposite in their role was that of Mrs Orange.

Group Orange has already been discussed for the clear disparity among students and lack of a sense of a community within the group. Mrs Orange began the activity with the introductory PowerPoint which seemed to set the lesson off to a promising start. Mrs Orange asked students questions about the subject matter which students were able to answer and the PowerPoint content also prompted questions from the students. However, the divide between Holly, James and the rest of the class emerged when students started the main activity. It was very apparent that Mrs Orange had not worked through the activity herself and was reliant on the ability of Holly and James. Though Mrs Orange told students to divide the supernova targets among them to combine into a master spreadsheet, she did not specifically assign targets to individual students. As there was no communication between the students, they did not arrange this themselves and Holly and James ended up working through all of the targets between the two of them.

The teacher was seen to come and go throughout the lesson, often leaving for up to ten minutes at a time. Though this was okay for Holly and James as they were more than capable of working through the activity, it was clear to the researcher and so should have been clear to the teacher that this was not the case for the other three students. Instead, they would have benefited from a greater input from Mrs Orange. Even when Mrs Orange did approach these students and noticed they were struggling, instead of offering help herself, she looked at what Holly and James were doing. This indicated that the

teacher had not been through the activity herself. It was judged that Mrs Orange presented behaviours of an absent facilitator. Though she introduced students to the subject matter and set them on the activity, she was absent for the majority of the lesson and rarely provided support to students when they needed it. Pianta *et al.* (2012) posit that negative student-teacher interactions inhibit development and engagement which seems to be what was seen in group Orange. Here the teacher appeared to be nonresponsive to the needs of the students, demotivating them and deterring their interest in the lesson task.

Where Mrs Yellow's advanced preparation of the activity and familiarity with her students was integral to her successful implementation, Mrs Orange's absence of preparation and interaction with her students was detrimental. These essentially bipolar implementations are prime examples of the influence of the teacher and their integral role, regardless of the classroom activity. When the researcher mentioned to Mrs Yellow that she had observed an implementation of the Hubble activity where the teacher had not been through it themselves first, she responded, "*yeah, see there's just no way that is going to work. You have to be prepared and know what you're doing*".

Pianta *et al.* (2012) emphasise that teacher sensitivity towards their students is crucial. Teachers should be aware of the needs of their students and react accordingly. Mrs Yellow for example recognised that the students would likely struggle with the scale of their graph given the large range in velocity values. In response she provided additional guidance for this part and students consequently did not lose engagement or report struggles with this aspect of the activity.

The teacher-student relationship in group Black was also of particular interest. Though the Hubble activity was implemented by the researcher, the teacher's role remained integral. Implementation in group Black was very much an informal process. The students were a small group and although of different ages, they were all described by the teacher to be of high ability. The students were very accepting of the researcher and it was very apparent that they held positive relationships with their teacher.

Throughout the duration of implementation, though the researcher delivered the activity and assisted with any problems or questions students had, the teacher would still wander among the students, seeing how they were getting on and asking them questions about their work. Though there were occasions where Mr Black took a step back to work, it was observed that as students became more animated in the lesson, so did he and this would consequently prompt him to interact with them.

It was heartening to see such positive relationships between the teacher and his students and the researcher often noted the exchange of stories and light-hearted jokes. Mr Black was also demonstrative of his sensitivity to the students and understanding of their needs, on several occasions he made suggestions on an extension of the activity that they could do based on what they had been covering in previous lessons. Mr Black suggested that Harry thought about what significant figures he

should use, and had the year 13 students add error bars to their measurements. These additional tasks, though small, effectively implemented differentiation across the group, even though the activity was delivered by the researcher. Pianta *et al.* (2012) emphasise how this kind of teacher behaviour encourages student interest and motivation and it is therefore no wonder these students demonstrated such high behavioural engagement (Fredricks *et al.*, 2004). It was also acknowledged that as the researcher was delivering the activity, it presented an opportunity for Mr Black to take a step back and remove himself from the lesson. However this was not the case and his presence and behaviour ensured that feelings of competence and relatedness were experienced by his students. As a result, positive teacher-student relationships were reinforced and their engagement was promoted as Mr Black fulfilled some of the students basic needs set out in self-determination theory (Gagne and Deci, 2005).

Group Yellow, Orange and Black were highly demonstrative of the integral role of the teacher, regardless of the OA. All students in these groups experienced the same activity but with varied implementation styles and varied relationships with their teacher, such variations in learning experiences have been highlighted in the preceding sub-chapters.

11.7 Overall Effectiveness

Having reviewed the effectiveness of the Hubble activity across each of the relevant groups through the lens of the evaluation framework, we again illustrate the overall effectiveness of the OAs via Reeves' (1994) technique in Figure 15, again each colour represents the relevant participant group.



Figure 15 – Illustration of the Overall Effectiveness of the Hubble's Expansion of the Universe OAs

Chapter 12

Discussion

This study began by identifying some of the key problems faced in England and Wales in terms of students' perceptions of their learning experiences in physics at secondary school. In response, the researcher sought to evaluate a series of educational OAs in order to identify specifically where positive learning experiences occurred, what constitutes an 'effective OA' and the implications for OA developers and implementing teachers.

The study followed a constructivist approach in methodology which also reflected the learning approach employed by the OAs. The research followed a multiple case study design and used mixed methods, specifically questionnaires, classroom observations, focus groups and interviews to gather and triangulate data. Quantitative data was statistically analysed using SPSS and thematic analysis was performed on qualitative data. All data was then combined and analysed through the lens of a specific evaluation framework comprised of seven dimensions in order to answer the four key research questions.

Having completed the data collection and analysis, this chapter now addresses each of the four research questions set out at the inception of this study.

12.1 Research Question 1

How are students' perceptions of positive learning experiences promoted within a physics lesson?

The first and foremost indicator of students' perceptions of their learning experiences, internal to the study rather than from pre-existing literature, was retrieved from the pre and post PPQ. Overall, where results were not accounted for by case, significant differences between students' perceptions of their learning experiences in their day-to-day physics lessons and in the implementation lesson were seen in three of the five scales, all of which went in favour for the post-PPQ. That is, for the scales Investigation Work, Confidence and General Interest, students reported significantly more positive scores in relation to the OA than to their day-to-day physics lessons. When broken down by case, significant differences were seen among three of the five scales for Down2Earth and two scales for Hubble's Expansion of the Universe.

For D2E, students reported significantly more positive perceptions towards the scales relating to Confidence, Perceived Relevance and General Interest with regards to the D2E activity than towards their day-to-day physics lessons. For the Hubble case, students reported significantly more positive perceptions towards Investigation Work and Confidence in the activity than in their day-to-day physics lessons.

However, the main argument for employing a mixed method approach to this study was to not only identify the *what* but also the *why*. Data collected through questionnaires, classroom observations, focus groups and interviews enabled the ability to do this. Chapter 2 reviewed the existing literature around students' perceptions and attitudes towards physics at school and a series of key influencing factors related to physics and science were identified, namely interest and enjoyment (Inter-alia, Hamlyn *et al.* 2017; Wellcome Trust, 2011; Sjøberg and Schreiner, 2010; Williams *et al.*, 2003), confidence (inter-alia, Gill and Bell, 2013; Tripney *et al.*, 2010; Barmby and Defty, 2006; Spall *et al.*, 2004), perceived relevance (inter-alia, Bevins *et al.*, 2005; Cleave, 2005; Osborne *et al.*, 2003), usefulness and career aspirations (inter-alia, Amettler and Ryder 2014; Archer *et al.* 2013), and classroom experiences (inter-alia, Rocard *et al.*, 2007; Lyons, 2006; Trumper, 2006a; Osborne and Collins, 2001). This study has gone a step further and identified what particular events and experiences must take place in a lesson in order to prompt positive perceptions from students in these areas.

This study has not only demonstrated a series of key factors that enhance students' learning experiences, but also that each one can be incorporated into a single lesson. Each of these factors are discussed in the following subsections.

12.1.1 Investigation and Exploration

Students' enjoyment of investigations and exploration activities was by no means a new contribution to the field as many studies have reported this reflection from students (e.g. Estyn, 2017; Ofsted, 2013; Tuan, 2005).

However, Chapter 3 emphasised that despite the recognised advantage of investigation and inquiry in the sciences, physics is largely still taught didactically (Siorenta and Jimoyiannis, 2008). Two of the key barriers to reform highlighted in Chapter 3 were related to issues of time constraints and cost or resource limitations meaning that teachers often resort to the safer, more straightforward didactic approach (InterAcademies Panel, 2010) and that some schools simply don't have the resources to implement practical, inquiry-based lessons (Songer *et al.*, 2003).

However, this study demonstrated that traditional, laboratory-based practical lessons are not the only means of implementing investigation and exploration within a physics lesson. The Down2Earth and Hubble activities revealed that students generally still felt they had completed an investigation in the

activity, with mean scores of 3.95 out of 5.00 overall, 4.01 for D2E and 3.88 for Hubble despite just using a computer, or in the case of groups Purple and Yellow in the Hubble case study, with just paper and calculators. Such activities were therefore implemented without the purchase of any equipment as the majority of schools now have access to computers, laptops or tablets (Coughlan, 2014). The OAs also did not demand that a significant proportion of lesson time was taken up by setting up and taking down equipment.

12.1.2 Experiences of Autonomy

One of the highly praised features of constructivism and its advantage over more traditional approaches is that it shifts the focus from a teacher-centred classroom (teaching) to a student-focused one (learning) (Ertmer and Newby, 2013). Student control, or more specifically, autonomy is considered by Gagne and Deci (2005) in their self-determination theory as one of three basic needs that must be fulfilled in order for students to feel intrinsically motivated. In terms relating specifically to physics education it has been reported by students as something that is often not fulfilled in their physics lessons (Angell *et al.*, 2004).

However, evidence from multiple methods indicated that in both cases, students were provided with opportunities to work autonomously and take control of their learning. Isabel from group Brown commented that she felt the teacher was in control of their learning in day-to-day lessons but felt the students were in control during the implementation lesson, thus highlighting a contrast in student autonomy. Multiple students across both cases also reflected in the post-PPQ that the freedom and independence they experienced was the element they enjoyed most about the activities. For example, *"we got to find things out independently", "we didn't have too much help – only when we needed it"* and *"the freedom of finding things out on our own"*.

However, despite these promising results, other feedback also demonstrated how restrictions and limitations in the OAs, particularly in the case of D2E, could impact negatively on students' learning experiences. Table 4 in Appendix 14 highlights a theme among what students liked least about the D2E activity was the restrictions of what parameters they could change on the simulator.

Evidence of both freedom and limitations from students demonstrated the importance of data source triangulation and the sequential explanatory design of the research that meant such discrepancies could be followed up and explored further. The researcher was able to discuss this with Mr Brown from the D2E case study in an interview who believed it was because in the Making a Meteorite activity, students were allocated groups and told which parameter they would be investigating. In response to the combined evidence, a suitable response was concluded in that the OAs in question will be redeveloped to include a discussion activity at the beginning that involves students thinking more about different approaches that could be used to investigate their research questions and relevant

scientific concepts. This, as reflected by Mr Brown would "give them [the students] a bit more ownership" over their learning.

12.1.3 Cooperation and Collaboration

This study demonstrated cooperative and collaborative learning to have a positive influence on students both academically and socially. Academically C/C opportunities allowed students to offer their skills in support of their peers and in turn receive support in other areas. This meant that students were able to support each other throughout the lesson and work was completed more efficiently. As described by Alfie in group Brown "*It's easier to work off each other, like some people might have stronger points in that subject so they can teach other people stuff*". Socially, such opportunities help to build a sense of community and relatedness among students, another of the three basic needs individuals require to feel intrinsically motivated (Gagne and Deci, 2005). This was particularly evident in group Purple where students came from a number of different schools. Students naturally gravitated into groups from their own school where they felt a sense of belonging and would also experience a greater feeling of confidence in their abilities than with a group of students they did not know or have any relationships with (Osterman, 2000). Such findings complement the argument posed by Dewey (1958) who claimed that it is through collaboration that learning occurs.

12.1.4 Novel, Unexpected Experiences

Having an element of novelty to the OA as suggested by Hidi (1990) did indeed prove to be advantageous in that it was seen to trigger students' situational interest. Generally, students would not expect to be able to simulate a meteor impact with Earth's surface, or to be able to analyse real data in order to calculate the age of the entire Universe, right from their classroom. Data from the two separate cases also demonstrated that novelty can be sourced from different areas, not just the context itself. For example evidence of the novelty of the context and the simulator itself in the Down2Earth case was demonstrated in students' unprompted comments in classroom observations, such as ones of disbelief, *"that crater is huge! I can't believe how big it is!"* and amazement, *"this is amazing, actually"*. Students in the Hubble case also yielded comments of disbelief but more related to not knowing certain tasks were possible, for example, *"this is great that you're actually able to use real data and use it to calculate the age of the Universe*". The novelty of working with real data was reported by 12 students in the Hubble case study as what they liked most about the activity.

12.1.5 Effective Differentiation for Building Confidence

Ofsted (2013) emphasise the vital feature of differentiation within both lesson plans and classroom conduct, and place it as the most effective method of raising achievement in science at school. However, Ofsted also highlight how not all teachers are embedding enough differentiation in their lessons.

Positive learning experiences are promoted when tasks are successfully differentiated so that they enhance students' confidence. Poor confidence is a commonly reported problem in physics education (e.g. Tripney *et al.* 2010; Spall *et al.*, 2004) and has been shown to be a reason for students' dislike of physics (Gill and Bell, 2013; Tripney *et al.*, 2010; Aschbacher and Roth, 2009).

Quantitatively and statistically speaking, confidence could be considered the greatest success of this initiative study. Both case studies demonstrated significantly higher levels of confidence among students towards the initiatives than towards their day-to-day physics lessons. Responses such as 'so I can do science', 'I feel like Einstein' and 'I should have done physics' were not only indicative of the huge influence appropriately targeted activities can have but also how powerful just one lesson can be. As emphasised early on in this thesis, the aim of this study was not to induce a change in students' attitudes, instead it sought to offer different, more positive learning experiences to students. However, the data that came to light was indicative of instances where students in fact reconsidered their perceptions. Such results are extremely promising given Chapter 2.5.2 highlighted that student confidence is much more problematic than actual ability (HOPE, 2016; Rodd et al., 2014) and that there are large correlations with students' confidence and their liking of a subject (Barmby and Defty, 2006; Williams et al., 2003). Also drawing back again to Gagne and Deci's (2005) self-determination theory, feelings of competence, which is considered to tie closely with confidence, is another one of the three basic needs required for students' intrinsic motivation. High confidence has therefore demonstrated various positive influences on students' learning experiences and this significant difference measured between students' day-to-day physics lessons and the initiative lesson should not be underestimated.

12.2 Research Question 2

How does teachers' planning and delivery of the Down2Earth and Hubble outreach activities influence students' learning experiences?

It is widely recognised that teachers are integral to students' learning experiences (Hamlyn *et al.*, 2017; Tripney *et al.*, 2010; Osborne and Collins, 2001) and the classroom climate, with traditional teaching approaches that primarily involve transmission of information from the teacher onto students triggering particularly unfavourable perceptions among students (Lyons, 2006). However, where the focus of the instructional design of the OAs here was shifted from teacher-centred learning to student-centred learning, it was initially unclear on the role that the teacher would play and how much influence they would have on the initiative lesson.

However, this study revealed that despite students displaying high levels of autonomy in the OAs, the teacher's role remained integral during implementation of the activities. Mr Black demonstrated the

huge influence a teacher could have over a lesson, even where the activity was delivered by the researcher. Mr Black's strong relationships with his students, demonstrated by knowing them individually and on a personal level, meant he still offered support throughout and even tailored small elements of the activity that incorporated effective differentiation into the activity. Mr Black's understanding of his students' individual needs demonstrated his positive relationship with them, a factor that is associated with student engagement (Roorda *et al.*, 2011; Pianta *et al.* 2012). Perhaps the most interesting point to be made here is that Mr Black was the only teacher who implemented this level of within-group differentiation, yet did not deliver the activity. Such within-group differentiation was also considered to promote students' sense of relatedness as set out by Gagne and Deci (2005) in their self-determination theory.

The most overwhelming evidence for the negative effects of poor preparation from the teacher and also the lack of student-teacher relationships was provided in group Orange in the Hubble case study. Not only was the teacher ill-equipped to support her students, having not gone through the activity herself prior to implementation, she also only demonstrated positive relationships with two students in the group. As a result, there was no sense of relatedness (Gagne and Deci, 2005) or community in the class and students, apart from the same two students, were neither cooperative nor collaborative in their learning. This was highly contrasting to observations of Mrs Yellow, who repeated her implementation over two years and whose increased confidence relayed onto implementation and her students' experiences, promoting more elements of collaborative learning and building the sense of community in the classroom.

12.3 Research Question 3

How does the implementation of the Down2Earth and Hubble outreach activities influence students' learning experiences?

This research question was from a development point of view and fed directly into Research Question 4 around the recommendations for future development and implementation of OAs. Opportunities are considered here to be events that promote positive learning experiences among students and in particular, some that may not be common in day-to-day physics lessons. Limitations of the OAs are recognised when students' learning experiences were impeded in some way, this could be either academically or socially.

The following sub sections highlight three opportunities and two threats that were identified throughout the study, from the implementation of the D2E and Hubble OAs and their comparison with day-to-day physics lessons.

12.3.1 Opportunity: Overcomes Barriers to Reform

Chapter 3.3 discussed various issues relating to why, despite the longstanding problems in science and physics education, many teachers are still resistant to the implementation of IBL. A lack of teacher specialists (Ward, 2018; National Audit Office, 2016), a lack of resources (Songer *et al.*, 2003) and time constraints (Wellcome Trust, 2016b), much of which is a result of the pressures of exam performances (Marsh, 2016), are commonly reported as reasons why inquiry-based or practical lessons are not implemented in more science lessons. Traditional, didactic lessons are seen to be easier to implement and of lower risk (Songer *et al.*, 2003) and therefore dominate physics education (Siorenta and Jimoyiannis, 2008).

However, this study has demonstrated two cases implemented across multiple groups whereby students have carried out investigations from start to finish within one lesson, covering compulsory content of the physics curriculum, at no cost with just the use of a computer or even pen and paper. Both the D2E and Hubble OAs are provided as 'Teacher Packages' so even where teachers are not physics specialists, all the material and information they need for implementation is readily available to them. This was demonstrated by Mr Brown (a chemistry specialist) in the results of his implementation.

The OAs provided in this study are freely available to all teachers and only require a computer with internet access, or in the case of Hubble, the activity can be done on paper. This therefore implies that the OAs developed here are a step towards overcoming this barrier. This also ties closely with the next point of time constraints. Time constraints are considered to be less of a limitation for the OAs of this study than typical practical, lab-based lessons. Where the latter require sufficient time to set up equipment and take it down, this is not a requirement for the D2E and Hubble activities.

In terms specifically relating to exam pressures, it is considered that the D2E and Hubble OAs are a good use of time as they not only cover compulsory curriculum content, they also address a number of scientific skills. The value of the Hubble OA in particular was emphasised by a comment from Mrs Yellow in her interview who said "*I think if they're doing physics A-level, I think every student should do this*".

12.3.2 Opportunity: An Unexpected, Novel Point of Engagement

This study has successfully demonstrated that a novel experience, such as simulating the consequences should a body from outer space impact with the Earth, or an unexpected experience, such as being able to work with real data to calculate the expansion rate and age of the Universe from the classroom, can successfully trigger students' situational interest. Given that situational interest is deemed to be particularly valuable with students who hold no individual interest in a subject (Hidi, 1990; Hidi and Harackiewicz, 2000), these OAs offer a prized opportunity to teachers who are faced with students who are consistently disengaged and uninterested in their physics lessons.

12.3.3 Opportunity: Outreach Activities Preventive of Entirely Didactic Teaching

This study has emphasised that even for OAs that focus on student-centred learning, the teacher's role remains integral and highly influential to implementation. However, one aspect that is ensured by the OAs is that although they accommodate for teachers' adjustments and personalisation of delivery, they are preventive of didactic delivery styles. Enough structure is embedded within the OAs so that teachers 'have' to implement them in such a way that is student-centred and allows for autonomous learning. The OAs would have to be dramatically transformed by a teacher, and would thus require a large amount of work and time in order to make them didactic in style.

12.3.4 Limitation: Differentiation Requires Improvement

Differentiation has been highlighted by Ofsted (2013) and Estyn (2017) as an area for improvement for many schools in England and Wales. In light of the findings, this also goes to be said for the OAs produced here.

Though in both cases students' reports of confidence was significantly higher on the post-PPQ, there were also reports that some students liked that the activity was challenging, others liked that it was simple, and some students reported that they disliked the difficulty of the activity. Such data is evidence for an inconsistency in the suitability of the activities for individual students. Feedback from Mr Brown also emphasised that further differentiation would be advantageous.

Though differentiation is recognised as a limitation of the OAs, this study successfully identified how it can be improved and developed in the future.

Firstly, the OAs will be provided in three versions, a top, middle and bottom. The top version is the most demanding and challenging to students and the bottom provides the most structure and guidance. Evidence suggests that differentiation was less problematic for the Hubble OAs than the D2E, it is considered that this is likely in part because all versions of the Hubble activity were provided to teachers, thus they were able to see the different levels. Exactly how the D2E activities could be differentiated was merely explained or suggested to teachers.

Secondly, the OAs will incorporate clear and progressive learning objectives. Clear so that students have an informed understanding of what they are aiming to achieve and progressive as this will facilitate differentiation.

12.3.5 Limitation: The Role of the Teacher

Discussion of the second research question of this study has dealt with this aspect in depth and so is not divulged into detail again here. However, it is considered that this integral role of the teacher does pose a limitation to the OAs and to those developed elsewhere. It is the belief of the researcher that regardless of how strong an OA may be, this only carries so far, and the success of implementation relies largely on how it is employed, interpreted and delivered by the teacher. This point is arguably the biggest limitation as it is largely out of the control of the OA designer. Even where detailed instructions or recommendations for implementation are provided, ultimately, all teachers will place their own spin on their delivery. Also, even more problematic is the teacher's social role in the classroom. This element is even more out of control of the OA designer, and teacher-student relationships are not something that can evolve overnight. Where the other limitations have been described and approaches to tackle them have been hypothesised, the same cannot be done for this point.

The researcher therefore emphasises this limitation to teachers. Though implementation of existing student-centred classroom OAs may suggest an opportunity to take a step back from the lesson, it should be emphasised that this is not the case. The teacher's role is still vital to the classroom environment and remains influential to students' senses of competence and belongingness (Gagne and Deci, 2005), teachers should stand as a facilitator to the students' learning throughout the lesson.

12.4 Research Question 4

What are the implications for the future development and utilisation of physics outreach activities for developers and teachers?

Some of the implications are clear from the descriptions of the OA advantages and limitations but are also summarised here. These implications will be taken forward by the researcher in future development of OAs and are recommended to other science OA developers and teachers in order to share experiences of best practice.

Ultimately, the researcher recommends that OA developers embed each of the five points described in Chapter 12.1, namely investigation and exploration; autonomous learning; cooperative and collaborative learning; novel, unexpected experiences; and effective differentiation. This study yielded substantial evidence that was demonstrative of their positive influence on students' learning experiences.

As described in the last point relating to research question 4 in terms of the teacher's role, regardless of the opportunities an educational OA might offer, this can still be overshadowed by a poorly prepared or disconnected teacher. Though practitioners should focus on how to best address students' learning experiences, they should also keep in mind the teacher, how they are going to implement and how to best support such implementation. Though not much can be done in terms of teacher-student relationships, the OA developer can, and should provide recommendations on points of discussion between the teacher and their students, and what teachers should be looking for and checking with their students. Teachers are also recommended that they should, at the very least, go through the OA

themselves ahead of implementation. Though these could impede on the barrier of time constraints emphasised by the Wellcome Trust (2016b), it is still considered to demand less time than designing a new lesson plan.

Practitioners should design OAs to have progressive objectives that are clearly and frequently presented to students and teachers (Quinn, 1996). Where they are progressive they also more readily encourage differentiation within the activity and break it down into smaller targets for students. More specifically related to differentiation, OA developers, where possible, should provide three versions of their activity, a top, middle and bottom, in order for the teacher to make an informed decision on which version best addresses the needs of his or her students. This recommendation is based largely on feedback from Mrs Yellow and Mr Brown in their interviews.

It is recommended that OAs should offer students opportunities for cooperative or collaborative learning. Such experiences have shown to enhance both students' academic and social experiences and can promote other classroom factors such as relatedness and competence, both in this study and existing literature (Gagne and Deci, 2005; Stipek, 2002; McMillan and Chavis, 1986). However, it is also emphasised that such opportunities should not be implemented at the expense of any vital learning experiences. That is, cooperation and collaboration should aid students' learning not limit it. All students should still experience each step of an activity as seen in group Yellow, rather than dividing steps among students as seen in group Purple which meant students did not grasp the entire process of the activity.

It is also recommended that OAs are provided as 'OA Packages' or 'Teacher Packages' that includes everything the teacher or students may need during implementation. That way all bases are covered and no additional work should be required from teachers to source anything they do not have. Such packages should include the following:

- A lesson plan with a breakdown of the activity, the lesson objectives, materials needed and time required
- An introductory PowerPoint presentation that sets the scene and introduces the subject matter
- Teacher Guides with guidelines for implementation, suggested areas of discussion, aspects students should consider and results they should be obtaining
- Student Worksheets for the students to work through and where they can record their results, and any notes they may wish to take
- Three differentiated version of the same activity in order for the teacher to best decide on an implementation strategy
- Hints and tips sheet this can be used to aid the students if they struggle at any point in the activity

• Background information – for the teacher in case the context of the OA is new and unfamiliar to them. This should also provide guidelines to any additional questions students may have

Having discussed how this study has addressed its defined research questions, the following final chapter concludes this study and sets out the specific contributions to knowledge.

Chapter 13

Conclusion

England, Wales and many other countries across the world are faced with a large population of students who are disengaged with physics at school. Due to its growing value and importance on the economy, a myriad of research has centred on students' attitudes and perceptions, and developing new initiatives and reforms.

Given that generally speaking, physics often appears to attract the most negative perceptions from students, and attracts the lowest level of applicants in post-16 education of the three sciences, physics was placed as the focus of this study. Existing literature around previous initiatives in science education lack thorough evaluation and transferable evidence that provides valuable insight to practitioners or teachers.

Evidence for students' disengagement with physics triggering in secondary education formed the audience for the outreach activities. Specific experiences of physics for secondary school students' that led to negative learning experiences provided areas to target.

Thorough literature reviews and preliminary thematic analysis of PPQ and observational data guided the construction of an evaluation framework used not only for the OAs in this study but that can be applied in other research into effectiveness of outreach activities or lessons. The framework comprised of dimensions that explored investigation and exploration, student autonomy, cooperation and collaboration among peers, novel or unexpected experiences and effective differentiation. Though there was evidence for areas in need of improvement, the OAs implemented in this study were demonstrative of the ability of incorporating each of these components into a single, curriculum-relevant lesson. The OAs are freely available, do not require any equipment other than possibly computers or tablets which are available in the majority of schools in the UK. The OAs themselves are provided as teacher packages so that all material is readily available and even non-physics specialist teachers can feel comfortable in implementing.

The specific contributions to knowledge of this study are therefore fourfold:

1. A robust evaluation framework

The evaluation framework that provided a lens through which the combined quantitative and qualitative data of this study was analysed, comprised of an amalgamation of original dimensions and ones from existing frameworks. Dimensions focused specifically on students' experiences with

qualitatively described dimensions and addresses the contextual elements of each implementation group in order to promote the transferability of findings. Though the framework offers sufficient detail into the research data and the inferences made, it also provides an overarching dimension (Overall Effectiveness) that refines and summarises such data, visually illustrating the strengths and weaknesses of the educational OAs.

As the framework focuses specifically on students' learning experiences and can account for contextual differences across classrooms, it can be applied to the evaluation of other educational OAs, whether developed externally or internally, computer or paper based. The framework is not confined to the components of a specific OA as the OA components, e.g. layout or structure, are not the focus. Instead the focus is placed on students' interaction with the OA and their experiences of its implementation.

The evaluation framework therefore provides a structure which can be transferred into other initiative evaluation studies in order to promote and understand interaction and learning experiences.

2. Astronomy as a context for physics in practice

Subchapter 4.1 emphasised the popularity of astronomy among students and how it has been described as *"the gateway science"* (Salimpour *et al.* 2018, p. 35; National Research Council, 2001, p. 4) and a *"point of engagement"* (Osborne and Collins, 2001, p. 457). This is due to the apparent interest in astronomy that is evident among wide student audiences.

However, though there is evidence for students' interest in astronomy and space more generally and the potential it offers as a point of engagement, there is an absence of research that specifically evaluates the application of astronomy as a context to the physics curriculum in terms of its influence on secondary school students' learning experiences. Although outreach projects with astronomy elements or enrichment activities exist, none were found to provide robust evaluation or focus specifically on contextualising the curriculum.

This study therefore went beyond simply asking students about what topics they find interesting. Instead it put this opportunity into practice and evaluated the implementation of astronomy-based OAs that tied directly to the English and Welsh physics curriculum and their influence on students' learning experiences. Findings from this study not only support the potential of astronomy as a point of engagement, but with the use of specific OA examples explicitly demonstrated how this context can be applied in practice to promote positive learning experiences in physics without straying from the curriculum.

3. Widely applicable recommendations for practitioners designing and evaluating school outreach activities

Subchapter 2.4 highlighted some of the initiatives in physics, science and STEM education that have been implemented in the UK in attempts to address students' disengagement. However, two of the key problems with such initiatives were the lack of in-depth evaluation of their impact or the narrow process of evaluation that was so specific to the particular initiative, it offered little insight and transferable recommendations for best practice to other practitioners seeking to develop educational OAs.

This study however, although implemented two specific OAs, the evaluation focused on students' interaction with the components of the OAs rather than the OAs themselves. As a result, an abundance of more general recommendations and implications resulted from findings that can be transferred to other initiatives. The specific OAs merely provided an example of such experiences and interactions that were observed.

The impact of this evaluation study therefore does not cease at the Down2Earth and Hubble OAs, instead it provides a basis and structure which practitioners can feed into their future OA developments. This ties back to the first point around the robust evaluation framework. Such a framework is therefore not limited to its use in evaluation of existing OAs but can also guide the design and development of new ones. Given that the framework specifies key components that can trigger positive learning experiences among students, practitioners should use this as a guide for structuring new OAs.

4. Initiatives over a single lesson can have an impact

A final contribution to knowledge is seen in the findings of this study that demonstrate how initiatives do not necessarily require large amounts of time or funding in order to have an impact. Though some researchers argue that only longitudinal initiatives can be effective in changing students' attitudes and perceptions, the researcher argues that this study has demonstrated that even enforcing alternative learning experiences in just a single lesson, can act as a stepping stone in the right direction, can encourage students' to reconsider their assumptions and cause significant differences in students' perceptions of their learning experiences. Examples of evidence of such claims includes student comments such as "so I can do science", "I feel like Einstein" and "I should have done physics".

The four key contributions to knowledge are considered to have implications at a micro, meso and macro level.

At the micro level, although the researcher is not expectant of a long term impact of this initiative, she is confident in the stepping stone it presents whereby if such OAs are fed into schemes and work, and the positive components of the OAs and recommendations are reinforced and sustained, then longer term changes could result. Furthermore if teachers outside of this study adopt such components and
translate them to their own classroom environment, they could also see evidence of students' alternative perceptions of learning experiences and take the first step towards such improvements.

At the meso level, findings are not considered to be exclusive to secondary physics education nor just to students in England and Wales. The findings and recommendations yielded in this study are deemed to be transferable to the development of OAs in other science subjects and across other countries. Lyons (2006) highlighted that students' in different countries appear to have similar experiences and perception of science education.

For OA developers, recommendations are provided in terms of how to construct classroom activities in such a way that makes them most readily available for implementation and how to consider the teacher's role in delivery.

At the macro level, this study is considered to be complementary to the accountability measures set out in England and Wales. Pressures faced by teachers and schools have been discussed in terms of exam performance, funding and time constraints and many outreach initiatives involve expenses and time for additional activities that are not necessarily directly related to the curriculum. However, this study has demonstrated how students' learning experiences in physics can be addressed and encouraged, without conflicting with such pressures, primarily stemming from government levels.

Given that reflexivity was embedded throughout this study and the researcher started the study as very much a novice in educational research, it is considered important to reflect on the researcher's own learning journey. Although students' learning experiences were very much the focus of the study, one of the researcher's most valuable insights was realising the importance of having an awareness of the needs of teachers. It became evident that in order to best support teachers, they need to be provided with off-the-shelf activities. If an activity can be implemented in multiple ways, teachers need to see each of these versions rather than be merely provided with suggestions. Teachers are the gatekeepers to the classroom and it is them who directly face accountability pressures. Therefore, in order to most benefit the students, practitioners must please the teachers. Teachers must buy in to the outreach activity, recognise its benefit to their students and feel confident that they can implement it.

Though this study is considered to have contributed to the wider field of physics education research, it is not without its limitations. On reflection, given the comprehensive, detailed nature of the research it would have been advantageous to have been able to track students throughout the implementation process. This was not done in this study as it is not straight forward and creates problems in ensuring anonymity of participants.

Observation data was the only data source that could be matched by occasion for individual participants, and this could not be matched with their questionnaire data, nor could their pre-PPQ be matched with their post-PPQ. Had this been possible, it would have permitted more powerful, paired

sample t-tests in the statistics rather than independent samples, and would also have enabled more thorough triangulation of data sources.

For example, where comparisons of observations of the behaviour of Lucy in group Blue in a normal lesson and in the implementation lesson revealed huge differences, such evidence would have been even more powerful had the researcher been able to identify Lucy's pre- and post-PPQ responses.

Additionally, the researcher also gathered demographic data in terms of gender and free-school-meal eligibility. These variables have only really been discussed in this study in terms of the representation of the sample, however much more could be investigated in terms of the outcome of implementation in relation to these specific variables. As the data collection was so vast, an exploration of differences across demographics extended beyond the scope of this study. However, as the researcher has this data readily available, such avenues shall be pursued for follow-up publications.

In terms of the next steps and future avenues beyond this study, there are two key opportunities. First and foremost is the implementation of the evaluation framework in other studies for other outreach activities. Though this study suggests the potential of the framework, it needs to be put into practice to further establish its validity, transferability and usefulness. Secondly, this study has demonstrated how a one-off implementation can successfully prompt students to reconsider their perceptions of physics and trigger situational interest. However the researcher recognises that this does not mean a long term change in attitudes and that in order for a more stable individual interest to develop, such experiences need to be fostered and sustained. A future study around how to sustain the responses from this study would therefore be beneficial.

References

Agrusti, G. 2013. Inquiry-based learning in Science Education: Why e-learning can make a difference, *Journal of e-Learning and Knowledge Society*, 9(2).

Ainley, M., Hidi, S. and Berndorff, D. 2002. Interest, Learning, and the Psychological Processes that Mediate their Relationship, *Journal of Educational Psychology* 94(3), pp. 545-561. DOI: 10.1037/0022-0663.94.3.545.

Albion, P. 1999. Heuristic Evaluation of Educational Multimedia: From Theory to Practice. *16th Annual Conference of the Australasian Society for Computers in Learning and Tertiary Education.* Available at: https://core.ac.uk/download/pdf/11039639.pdf [Accessed: 09 August 2018].

Alderson, P. 2005. Designing Ethical Research with Children. In: Farrell, A. (Ed). *Exploring Ethical Research with children*. Buckingham: Open University Press, pp. 27-36.

Aldridge, J. M. Fraser, B., Bell, L. and Dorman, J. 2012. Using a New Learning Environment Questionnaire for Reflection in Teacher Action Research, *Journal of Science Teacher Education* 23(3), pp. 259-290. DOI: 10.1007/s10972-012-9268-1

Alvesson, M. and Sköldberg. 2009. *Reflexive Methodology: New Vistas for Qualitative Research*. 2nd ed. London: SAGE Publications.

Ametller, J. and Ryder, J. 2014. The Impact of Science Curriculum Content on Students' Subject Choices in Post-Compulsory Schooling. In: Henriksen, E., Dillon, J. and Ryder, J. (Eds). *Understanding Student Participation and Choice in Science and Technology Education*, pp. 103-118. London: Springer.

Anderman, E. 2002. School Effects on Psychological Outcomes during Adolescence. *Journal of Educational* Psychology 94(4), pp. 795-809. DOI: 10.1037/0022-0663.94.4.795

Angell, C., Guttersrud, Ø., Henriksen, E. and Isnes, A. 2004. Physics: Frightful, but Fun. Pupils' and Teachers' Views of Physics and Physics Teaching, *Science Education* 88(5), pp. 683-706. DOI: 10.1002/sce.10141

Archer, L., Osborne, J., DeWitt, J., Dillon, J. and Wong, B. 2013. *ASPIRES: Young People's Science and Career Aspirations, Age 10-14*. Available at: https://www.kcl.ac.uk/sspp/departments/education/research/ASPIRES/ASPIRES-final-report-December-2013.pdf [Accessed: 14 October 2018].

Arslan, A. 2014. Transition between Open and Guided Inquiry Instruction, *Procedia – Social and Behavioral Sciences* 141, pp. 407-412. DOI: 10.1016/j.sbspro.2014.05.071

Arthur, S. and Nazroo, J. 2003. Designing Fieldwork Strategies and Materials. In: Ritchie, J. and Lewis, L. eds. *Qualitative research practice: A guide for social science students and researchers*. London: SAGE Publications, pp. 109-137.

Aschbacher, P. and Roth, E. 2009. Is Science Me? High School Students' Identities, Participation and Aspirations in Science, Engineering and Medicine, *Journal of Research in Science Teaching* 47(5), pp. 564-582. DOI: 10.1002/tea.20353

Aslam, F., Adefila, A. and Bagiya, Y. 2018. STEM Outreach Activities: an Approach to Teachers' Professional Development, *Journal of Education for Teaching International Research and Pedagogy* 44(1), pp. 58-70. DOI: 10.1080/02607476.2018.1422618

Astronomy and Geophysics. 2008. High Impact Astronomy. December, 49(6), p. 6.4. DOI: 10.1111/j.1468-4004.2008.49604_1.x

Ausubel, D. 1963. The psychology of meaningful verbal learning. Oxford, England: Grune & Stratton.

Bailey, G. 2014. Accountability and the Rise of "Play Safe" Pedagogical Practices, *Education & Training* 56(7), pp. 663-674. DOI: 10.1108/ET-07-2014-0081

Bailey, J. and Slater, T. 2005. Finding the Forest Amid the Trees: Tools for Evaluating Astronomy Education and Public Outreach Projects, *Astronomy Education Review*, 3(2). DOI: 10.3847/AER2004016

Banchi, H. and Bell, R. 2008. 'The Many Levels of Inquiry', *Science and Children*, 46(2) pp. 26-29. Available at: http://www.miseagrant.umich.edu/lessons/files/2013/05/The-Many-Levels-of-Inquiry-NSTA-article.pdf [Accessed: 03 May 2015].

Bandura, A. 1977. Confidence: Toward a unifying theory of behavioral change. *Psychological Review* 84(2), 191-215. DOI: 10.1037//0033-295x.84.2.191

Barmby, P. and Defty, N. 2006. Secondary School Pupils' Perceptions of Physics, Research in *Science & Technological Education* 24(2), pp. 199-215. DOI: 10.1080/02635140600811585

Barmby, P., Kind, P. and Jones, K. 2008. Examining Changing Attitudes in Secondary School Science, *International Journal of Science Education*, 30(8), pp. 1075-1093. DOI: 10.1080/09500690701344966

Bartlett, S. 2017. *New Physics Educational Resources*. Blogs: *Physics and Astronomy Outreach* 16 January Available at: http://blogs.cardiff.ac.uk/physicsoutreach/2017/01/16/new-physics-educational-resources/ [Accessed: 22 November 2018]

Bartlett, S. 2018. Hubble Plots in High School. In: Fitzgerald, M., James, M., Buxner, C.R. and White, S. Eds. *Robotic Telescopes, Student Research and Education (RTSRE)*. San Diego, US, 18-21 June, 2017, pp. 261-269. DOI: 10.32374/rtsre.2017.024

Bartlett, S., Fitzgerald, M., McKinnon, D., Danaia, L. and Lazendic-Galloway, J. 2018. Astronomy and Science Student Attitudes (ASSA): A Short Review and Validation of a New Instrument, *Journal of Astronomy and Earth Sciences Education* 5(1). DOI: 10.19030/jaese.v5i1.10190.

Baruch, J. 2015. A robotic telescope for science and education. *Astronomy and Geophysics* 56(2), pp. 2.18–2.21. DOI: 10.1093/astrogeos/atv061

Bassey, M. 1981. Pedagogic Research: On the Relative Merits of Search for Generalisation and Study of Single Events, *Oxford Review of Education* 7(1), pp. 73-94. DOI: 10.1080/0305498810070108

Beare, R. 2007. Investigation into the Potential of Investigative Projects Involving Powerful Robotic Telescopes to Inspire Interest in Science, *International Journal of Science Education* 29(3), pp. 279-306. DOI: 10.1080/09500690600620938.

Bell, D. 2014. *The Perceived Success of Interventions in Science Education*. Available at: https://wellcome.ac.uk/sites/default/files/wtp056459 0.pdf [Accessed: 14 October 2018].

Bell, R. L., Smetana, L. and Binns, I. 2005. Simplifying inquiry instruction, *The Science Teacher*, 72(7), pp. 30-33.

Bernard, Z. 2017. Here's How Technology is Shaping the Future of Education. *Business Insider*. Available at: http://uk.businessinsider.com/how-technology-is-shaping-the-future-of-education-2017-12?r=US&IR=T [Accessed: 17 October 2018].

Bevins, S., Brodie, M. and Brodie, E. 2005. A study of UK secondary school students' perceptions of science and engineering, *European Educational Research Association Annual Conference*, Dublin, 7-10 September.

Birch, S. and Ladd, G. 1997. The Teacher-Child Relationship and Children's Early School Adjustment, *Journal of School Psychology* 35(1), pp. 61-79. DOI: 10.1016/S0022-4405(96)00029-5

Braun, V. and Clarke, V. 2006. Using Thematic Analysis in Psychology, *Qualitative Research in Psychology* 3(2), pp. 77-101. DOI: 10.1191/1478088706qp063oq

Brill, R., Grayson, H., Kuhn, L. and O'Donnell, S. 2018. *What Impact Does Accountability Have on Curriculum, Standards and Engagement in Education?* Available at: https://www.nfer.ac.uk/media/3032/nfer_accountability_literature_review_2018.pdf [Accessed: 09 December 2018]

British Educational Research Association [BERA]. 2018. *Ethical Guidelines for Educational Research* (4th Edn) London: BERA. Available at: https://www.bera.ac.uk/wp-content/uploads/2018/06/BERA-Ethical-Guidelines-for-Educational-Research 4thEdn 2018.pdf?noredirect=1 [Accessed: 26 November 2018]

Broughton, N. 2013. *In the Balance: The STEM Human Capital Crunch*. Available at: http://www.smf.co.uk/wp-content/uploads/2013/03/Publication-In-The-Balance-The-STEM-human-capital-crunch.pdf [Accessed: 14 October 2018].

Bruner, J.S. 1986. Actual Minds, Possible Worlds. London: Harvard University Press.

Burrell, G. and Morgan, G. 1979. *Sociological Paradigms and Organisational Analysis*. London, UK: Heinemann Educational Books.

Cakir, M. 2008. Constructivist Approaches to Learning in Science and Their Implications for Science Pedagogy: A Literature Review, *International Journal of Environmental and Science Education*, 3(4), pp. 193-206.

Campbell, D. 1957. Factors Relevant to the Validity of Experiments in Social Settings, *Psychological Bulletin*, 54(4), pp. 297-312. DOI: 10.1037/h0040950

Campbell, D. and Fiske, D. 1959. Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix, *Psychological Bulletin*, 56(2), pp. 81-105. DOI: 10.1037/h0046016

Cannell, C., Oksenberg, L. and Fowler, F. 1989. *New Techniques for Pretesting Survey Questions. Available at:*

https://www.psc.isr.umich.edu/dis/infoserv/isrpub/pdf/NewTechniquesPretestingSurveyQuestions_O CR.pdf [Accessed: 15 October 2018].

CBI/Pearson. 2015. *Education and Skills Survey 2015*. Available at: http://news.cbi.org.uk/reports/education-and-skills-survey-2015/ [Accessed: 15 July 2015].

CBI/Pearson. 2017. *Helping the UK Thrive: CBI/Pearson Education and Skills Survey 2017*. Available at: http://www.cbi.org.uk/index.cfm/_api/render/file/?method=inline&fileID=DB1A9FE5-5459-4AA2-8B44798DD5B15E77 [Accessed: 14 October 2018].

Centra, J.A. 1973. Self-Ratings of College Teachers: A Comparison with Student Ratings, *Journal of Educational* Measurement 10(4), pp. 286-295. DOI: 10.1111/j.1745-3984.1973.tb00806.x

Chatterjee, S., Williamson, V., McCann, K. and Peck, L. 2009. Surveying Students' Attitudes and Perceptions toward Guided-Inquiry and Open-Inquiry Laboratories, *Journal of Chemical Education* 86(12), pp. 1427-1432 DOI: 10.1021/ed086p1427

Chen, X. 1997. *Students' Peer Groups in High School: The Pattern and Relationships to Educational Outcomes.* Available at: https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=97055 [Accessed: 15 October 2018].

Chubb, K., Hood, R., Wilson, T., Holdship, J. and Hutton, S. 2017. Discovering New Variable Stars at Key Stage 3. *Physics Education* 52(3). Available at: http://iopscience.iop.org/article/10.1088/1361-6552/aa649f/meta [Accessed: 02 November 2017].

Clark, G. 1998. Telescopes in Education (TIE): Bringing Live Astronomy to the Classroom Utilizing Remotely Controlled Telescopes. In: Okyudo, M., Ebisuzaki, T. and Nakayama, M. Eds. *Astronomical Education with the Internet*. Tokyo: Universal Academic Press, pp. 37-44.

Clark, L. and Watson, D. 1995. Constructing Validity: Basic Issues in Objective Scale Development. *Psychological Assessment* 7(3), pp. 309-319. DOI: 10.1037/1040-3590.7.3.309

Cleaves, A. 2005. The formation of science choices in secondary school', *International Journal of Science Education*, 27(4), pp. 471-486.

Clements, D.H. and Battista, M. T. 1990. Constructivist Learning and Teaching, *The Arithmetic Teacher* 38(1), pp. 34-35.

Coe, R., Aloisi, C., Higgins, S. and Major, L.E. 2014. *What Makes Great Teaching?: Review of the Underpinning Research*. Available at: https://www.suttontrust.com/wp-content/uploads/2014/10/What-Makes-Great-Teaching-REPORT.pdf [Accessed: 09 December 2018]

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd Edition. New Jersey, US: Lawrence Earlbaum Associates.

Cohen, L., Manion, L., Morrison, K. 2011. Research Methods in Education. 7th Edition. Abingdon, UK: Routledge.

Collins, G., Melosh, H. and Marcus, R. 2005. Earth Impact Effects Program: A Web-Based Computer Program for Calculating the Regional Environmental Consequences of a Meteoroid Impact on Earth, *Meteoritics and Planetary Science* 40(6), pp. 817-840.

Cooper, P. 1993. Paradigm Shifts in Designed Instruction: From Behaviourism to Cognitivism to Constructivism, *Educational Technology* 33(5), pp. 12-19. Available at: https://www.jstor.org/stable/44428049?seq=1#page_scan_tab_contents [Accessed: 25 September 2018].

Corter, J. E., Nickerson, J.V., Esche, S. and Chassapis, C. 2004. Remote versus hands-on labs: a comparative study, *Frontiers in Education 2004.* 34th Annual. Savannah, 20-23 October. DOI: 10.1109/FIE.2004.1408586

Corter, J., Nickerson, J., Esche, S., Chassapis, C., Im, S. and Ma, J. 2007. Constructing Reality: A Study of Remote, Hands-On, and Simulated Laboratories, *ACM Transactions on Computer-Human Interaction (TOCHI)* 14(2). DOI: 10.1145/1275511.1275513.

Coughlan, S. 2014. *Tablet Computers in '70% of Schools'*. Available at: https://www.bbc.co.uk/news/education-30216408 [Accessed: 15 October 2018].

Coward, D.M., Heary, A., Venville, G., Todd, M., Laas-Bourez, M., Zadnik, M., Klotz, A., Boer, M. and Longnecker, N. 2011. The Zadko telescope: A resource for science education enrichment, *Advances in Space Research*, 47(11), pp. 1922-1930. DOI: 10.1016/j.asr.2011.01.005

Crawford, B. 2007. Learning to Teach Science as Inquiry in the Rough and Tumble of Practice, *Journal of Research in Science Teaching*, 44(4), pp. 613-642.

Creswell, J. 2002. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 2nd Ed. Thousand Oaks, US: SAGE Publications.

Creswell, J. 2009. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 3rd Ed. Thousand Oaks, US: SAGE Publications.

Creswell, J. and Plano-Clark, V. 2011. *Designing and Conducting Mixed Methods Research*. 2nd Ed. Los Angeles, US: SAGE Publications.

Creswell, J. Plano-Clark, V., Gutmann, M. and Hanson, W. 2003. Advanced Mixed Method Research Designs. In: Tashakkori, A. and Teddlie, C. Eds. *Handbook of Mixed Methods in Social and Behavioural Research*. Thousand Oaks, US: SAGE Publications, pp. 209-240.

Creswell, J. and Miller, D. 2000. Determining Validity in Qualitative Inquiry, Theory Into Practice 39(3), pp. 124-130. DOI: 10.1207/s15430421tip3903_2

Danaia, L., Mckinnon, D., Parker, Q., Fitzgerald, M. and Stenning, P. 2012. Space to Grow: LOCGT.net and Improving Science Engagement in Schools, *Astronomy Education Review*, 11(1). DOI: 10.3847/AER2012007

Deci, E. and Ryan, R. 2012. Motivation, Personality, and Development within Embedded Social Contexts: An Overview of Self-Determination Theory. In: Ryan, R. Ed. *The Oxford Handbook of Human Motivation*. New York, US: Oxford University Press, pp. 85-110.

de Jong, T. 2006. Technological Advances in Inquiry Learning, *Science*, 312(5773), pp. 532-533. *Science in the Classroom*. Available at: http://portal.scienceintheclassroom.org/ [Accessed: 13 May 2015].

de Jong, T. 2013. Deliverable D1.1 Go-Lab learning spaces specification, *Go-Lab Deliverables. HAL*. Available at: https://telearn.archives-ouvertes.fr/hal-00979335 [Accessed: 13 May 2015].

Deloitte. 2012. *The Importance of Physics to the UK Economy*. Available at: http://www.iop.org/publications/iop/2012/file 58713.pdf [Accessed: 14 October 2018].

den Brok, P., Bergen, T. and Brekelmans, M. 2006. 'Convergence and Divergence between Students' and Teachers' Perceptions of Instructional Behaviour in Dutch Secondary Education'. In: Fisher, D. and Khine, M. Eds. *Contemporary Approaches to Research on Learning Environments*. Singapore: World Scientific, pp. 125-160.

Denscombe, M. 2008. Communities of Practice: A Research Paradigm for the Mixed Methods Approach, *Journal of Mixed Methods Research*, 2(3), pp. 270-283. DOI: 10.1177/1558689808316807

Denzin, N. 1971. The Logic of Naturalistic Inquiry, *Social Forces* 50(2), pp. 166-182. DOI: 10.2307/2576935

Denzin, N. 1978. *The research act: a theoretical introduction to sociological methods*. 2nd ed. New York: McGraw-Hill.

Denzin, N. 1984. The Research Act. Englewood Cliffs, US: Prentice Hall.

Department for Education. 2014. *National curriculum in England: science programmes of study – key stage 4.* . Available at:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/381380/Science_KS4_PoS_7_November_2014.pdf [Accessed: 26 January 2016].

Department for Education. 2015. *National Curriculum in England: Science Programmes of Study*. . Available at: https://www.gov.uk/government/publications/national-curriculum-in-england-science-

programmes-of-study/national-curriculum-in-england-science-programmes-of-study [Accessed 14 October 2018].

Department for Education. 2017a. *Policy Paper: English Baccalaureate (EBacc)*. Available at: https://www.gov.uk/government/publications/english-baccalaureate-ebacc/english-baccalaureate-ebacc [Accessed: 19 November 2018]

Department for Education/CooperGibson Research. 2017b. Understanding Schools' Responses to the Progress 8 Accountability Measure. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/629 517/School response to Progress 8 Report.pdf [Accessed: 19 November 2018]

Department for Education 2018. Secondary Accountability Measures: Guide for Maintained Secondary Schools, Academies and Free Schools. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/748 497/Secondary_accountability_measures_guidance__-October_2018.pdf [Accessed: 19 November 2018].

DeVries, R. 2002. *What Does Research on Constructivist Education tell us about Effective Schooling?* Available at: http://www2.education.uiowa.edu/html/iae/iae-z-op-devries-1-5.pdf [Accessed: 22 October 2018]

DeWalt, K. and DeWalt, B. 2002. *Participant Observation: A Guide for Fieldworkers*. Walnut Creek, US: AltaMira Press.

Dewey, J. 1913. Interest and Effort in Education. Boston, US: Houghton Mifflin Company.

Dewey, J. 1958. Experience and Education. New York, US: Macmillan

Dweck, C. 1986. Motivational Processes Affecting Learning, *American Psychologist* 41(10), pp. 1040-1048.

Dweck, C. 2016. Mindset: The New Psychology of Success. New York, US: Ballantine Books.

Edelson, D.C., Gordin, D.N. and Pea, R.D. 1999. Addressing the Challenges of Inquiry-Based Learning through Technology and Curriculum Design. Journal of the Learning Sciences, *Journal of the Learning Sciences* 8(3-4), pp. 391-450. DOI: 10.1080/10508406.1999.9672075

Education Achievement Service (EAS). No Date. *About the Education Achievement Service for South East Wales*. Available at: https://sewales.org.uk/ [Accessed: 15 October 2018].

Education Endowment Foundation. 2018a. *Let's Think Secondary Science*. Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/lets-think-secondary-science/ [Accessed: 11 December 2018]

Education Endowment Foundation. 2018b. *Spaced Learning*. Available at: https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/spaced-learning/ [Accessed: 11 December 2018]

Eistenhardt, K. and Graebner, M. 2007. Theory Building from Cases: Opportunities and Challenges, *Academy of Management Journal* 50(1). DOI: 10.5465/amj.2007.24160888

Elo, S. and Kyngäs, H. 2008. The Qualitative Content Analysis Process. *Journal of Advanced Nursing* 62(1), pp. 107-115. DOI: 10.1111/j.1365-2648.2007.04569.x

English, H.B. and English, A.C. 1958. *A Comprehensive Dictionary of Psychological and Psychoanalytical Terms: A Guide to Usage*. Oxford, UK: Longmans Green.

English Oxford Dictionary. 2018. Available at: https://en.oxforddictionaries.com/ [Accessed: 11 December 2018]

Erickson, F. 2012. Qualitative Research Methods for Science Education', in Fraser, B., Tobin, K. and Campbell, M. (Eds.) *Second International Handbook of Education*. Netherlands: Springer, pp. 1451-1469.

Ertmer, P. and Newby, T. 2013. Behaviorism, Cognitivism, Constructivism: Comparing Critical Features from an Instructional Design Perspective, *Performance Improvement Quarterly* 26(2), pp. 43-71. DOI: 10.1002/piq.21143

Estyn. 2017. *Science at Key Stage 3 and Key Stage 4*. Available at: https://www.estyn.gov.wales/sites/default/files/files/documents/Science%20at%20key%20stage%203 %20and%20key%20stage%204%20%28002%29.pdf [Accessed: 22 October 2018]

Fensham, P.J. 2009. Real World Contexts in PISA Science: Implications for Context-Based Science Education, *Journal of Research in Science Teaching* 46(8), pp. 884-896. DOI: 10.1002/tea.20334

Finch, H. and Lewis, J. 2003. Focus Groups. In: Ritchie, J. and Lewis, L. eds. *Qualitative research practice: A guide for social science students and researchers*. London: SAGE Publications, pp. 170–198.

Firestone, W. 1987. Meaning in Method: The Rhetoric of Quantitative and Qualitative Research, *Educational Researcher*, 16(7), pp. 16-21. DOI: 10.2307/1174685

Fisher, D. and Fraser, B. 1983. A Comparison of Actual and Preferred Classroom Environments as Perceived by Science Teachers and Students, *Journal of Research in Science Teaching 20*(1). DOI: 10.1002/tea.3660200106

Fitzgerald, M. 2015a. Design Principles, Implementation and Evaluation for Inquiry-Based Astronomy: An Investigation of the Issues Surrounding Sufficient Teacher Professional Development in Large-Scale Astronomical Initiatives. PhD thesis. Macquarie University.

Fitzgerald, M. 2015b. Skype Conversation with Sophie Bartlett, 13 August.

Fitzgerald, M., Hollow, R., Rebull, I. M., Danaia, L. and McKinnon, D. 2014. A Review of High School Level Astronomy Student Research Projects Over the Last Two Decades, *Publications of the Astronomical Society of Australia*, 31. DOI: 10.1017/pasa.2014.30

Fitzgerald, M., McKinnon, D., Danaia, L. and Deehan, J. 2015a. Large-Scale Inquiry-Based Astronomy Intervention Project: Impact on Students' Content Knowledge Performance and Views of their High School Science Classroom, *Research in Science Education* 46(6), pp. 901-916. DOI: 10.1007/s11165-015-9486-6

Fitzgerald, M., McKinnon, D., Danaia, L. and Bartlett, S. In Writing. Differences in Perception of High School Science: Students' and Teachers' Views.

Fitzgerald, M., McKinnon, D. and Danaia, L. 2015b. Inquiry-Based Educational Design for Large-Scale High School Astronomy Projects Using Real Telescopes, *Journal of Science Education and Technology*. DOI: 10.1007/s10956-015-9560-x

Fraser, B. 1982. Differences between Student and Teacher Perceptions of Actual and Preferred Classroom Learning Environment, *Educational Evaluation and Policy Analysis* 4(4), pp. 511-519. DOI: 10.2307/1163661

Fredricks, J., Blumenfeld, P. and Paris, A. 2004. School Engagement: Potential of the Concept, State of the Evidence, *Review of Educational Research* 74(1). DOI: 10.3102/00346543074001059

Furrer, C., Skinner, E. and Pitzer, J. 2014. The Influence of Teacher and Peer Relationships on Students' Classroom Engagement and Everyday Motivational Resilience, *Teachers College Record* 116(13), pp. 101-123.

Gagne, M. and Deci, E. 2005. Self-Determination Theory and Work Motivation, *Journal of Organizational Behaviour*, 26, pp. 331-362. DOI: 10.1002/job.322

Galda, P. 2017. Revisiting Bias in Qualitative Research: Reflections on Its Relationship with Funding and Impact, *International Journal of Qualitative Methods* 16(1). DOI: 10.1177/1609406917748992

Gall, M., Borg, W. and Gall, J. 1996. *Educational Research: An Introduction*. 6th ed. New York: Longman Publishers.

Galton, M. 2009. Moving to secondary school: Initial encounters and their effects, *Perspectives on Education*, 2, pp. 5-21.

Gardner, P. L. 1975. Attitudes to Science: A Review, *Studies in Science Education*, 2(1), pp. 1-41. DOI: 10.1080/03057267508559818

George, M. 2017. *Shortage of Specialist Science Teachers 'is Biggest Challenge Facing Schools'*. Available at: https://www.tes.com/news/shortage-specialist-science-teachers-biggest-challenge-facing-schools [Accessed: 22 October 2018]

George, D. and Mallery, M. 2010. *SPSS for Windows Step by Step: A Simple Guide and Reference*, (17.0 Update). Boston: Pearson.

Gill, T. and Bell, J. 2013. What Factors Determine the Uptake of A-Level Physics? *International Journal of Science Education* 35(5), pp. 753-772. DOI: 10.1080/09500693.2011.577843

Gold, R. 1958. Roles in Sociological Field Observations. *Social Forces*, 36(3), pp. 217-223. DOI: 10.2307/2573808

Gomez, E. and Fitzgerald, F. 2017. Robotic Telescopes in Education, *Astronomical Review* 1(13). DOI: 10.1080/21672857.

Gonzales, P., Guzman, J.C., Jocelyn, L., Katsberg, D. and Williams, T. 2004. *Highlights From the Trends in International Mathematics and Science Study: TIMMS 2003.* Washington, DC: U.S. Department of Education, National Center for Education Statistics. Available at: https://files.eric.ed.gov/fulltext/ED483080.pdf [Accessed: 10 December 2018]

Gooday, G., Lynch, J., Wilson, K. and Barsky, C. 2008. *Does Science Education Need the History of Science?* Isis 99(2), pp. 322-330. DOI: 10.1086/588690

Goodrum, D., Hackling, M. and Rennie, L. 2001. *The Status and Quality of Teaching and Learning of Science in Australian Schools*. Research Report for the Department of Education, Training and Youth Affairs. Canberra, Australia: DETYA.

Gould, R., Dussault, M. and Sadler, P. 2007. What's Educational about Online Telescopes?: Evaluating 10 Years of MicroObservatory. *Astronomy Education Review* 5(2), pp. 127-145. DOI: 10.3847/AER2006022

GOV.UK. 2017. Find and Compare Schools in England. Available at: https://www.compare-school-performance.service.gov.uk/ [Accessed: 11 October 2018]

Grant, A. 1995. The Effective use of Laboratories in Undergraduate Courses. *International Journal of Mechanical Engineering Education* 23(2), pp. 95-101.

Grant, A., Seymour, K. and Smith, J. 2017. *Evaluation of the Lab in a Lorry Programme*. Available at: http://www.iop.org/resources/lab-in-a-lorry/evaluation-2005/file_71467.pdf [Accessed: 14 October 2018].

Grasha, A. 1994. A Matter of Style: The Teacher as Expert, Formal Authority, Personal Model, Facilitator, and Delegator, *College Teaching* 42(4), pp. 142-149. DOI: 10.1080/87567555.1994.9926845

Greene, J. 2008. Is Mixed Methods Social Inquiry a Distinctive Methodology?. *Journal of Mixed Methods Research*, 2(1), pp. 7-22. DOI: 10.1177/1558689807309969

Greene, J., Caracelli, V. and Graham, W. 1989. Toward a Conceptual Framework for Mixed-Method Evaluation Designs, *Educational Evaluation and Policy Analysis* 11(3), pp. 255-274. DOI: 10.2307/1163620

Guba, E. and Lincoln, Y. 1981. Effective Evaluation: Improving the Usefulness of Evaluation Results through Responsive and Naturalistic Approaches. San Francisco, US: Jossey-Bass.

Guthrie, J. and Wigfield, A. 2000. Engagement and Motivation in Reading. In: Kamil, M., Mosenthal, P., Pearson, D. and Barr, R. *Handbook of Reading Research*, Volume 3. New York, US: Routledge, pp. 403-422.

Hadjerrouit, S. 2010. A Conceptual Framework for Using and Evaluating Web-Based Learning Resources in School Education. *Journal of Information Technology Education*, 9, pp. 53-79.

Hamlyn, R., Matthews, P. and Shanahan, M. 2017. *Young People's Views on Science Education*. Available at: https://wellcome.ac.uk/sites/default/files/science-education-tracker-report-feb17.pdf [Accessed: 14 October 2018].

Haury, D.L. 1993. Teaching Science through Inquiry. ERIC/CSMEE Digest. Available at: https://files.eric.ed.gov/fulltext/ED359048.pdf [Accessed: 16 November 2017]

Herron, M. D. 1971. The nature of scientific inquiry, *School Review*, 79(2), pp. 171-212. DOI: 101086/442968

Hess, J. 1968. Group Interviewing. In: King, R.L. Ed. *New Science of Planning*. Chicago, US: American Marketing Association, pp. 51-84.

Hidi, S. 1990. Interest and its Contribution as a Mental Resource for Learning, Review of Educational Research 60(4), pp. 549-571. DOI: 10.3102/00346543060004549

Hidi, S. and Harackiewicz, J. 2000. Motivating the Academically Unmotivated: A Critical Issue for the 21st Century, *Review of Educational Research* 7(20). DOI: 10.3102/00346543070002151.

Hidi, S. and Renninger, K. 2006. The Four-Phase Model of Interest Development, *Educational Psychologist* 41(2), pp. 111-127. DOI: 10.1207/s15326985ep4102 4

Hitchcock, G. and Hughes, D. 1995. *Research and the Teacher: A Qualitative Introduction to Schoolbased Research*. 2nd Edition. London: Routledge

Hinkin, T., Tracey, B. and Enz, C. 1997. Scale Construction: Developing Reliable and Valid Measurement Instruments, *Journal of Hospitality and Tourism Research* 21(1). DOI: 10.1177/109634809702100108

Hook, C.M. and Rosenshine, B.V. 1979. Accuracy of Teacher Reports of Their Classroom Behaviour. *Review of Educational Research* 49(1), pp. 1-11. DOI: 10.3102/00346543049001001

Horizons in Physics Education (HOPE). 2016. *Horizons in Physics Education: Final Report*. Available at: https://hopenetwork.eu/sites/default/files/2013_3710%20_FR_%20HOPE_pub.pdf [Accessed: 14 October 2018].

Hung, W. 2011. Theory to reality: A few issues in implementing problem-based learning, *Educational Technology Research and Development*, 59(4), pp. 529-552. DOI: 10.1007/s11423-011-9198-1

Hunt, S., Sparkman, R. and Wilcox, J. 1982. The Pretest in Survey Research: Issues and Preliminary Findings, *Journal of Marketing Research*, 19(2), pp. 269-273. DOI: 10.2307/3151627

Huppert, J., Lomask, M. and Lazarowitz, R. 2002. Computer Simulations in the High School: Students' Cognitive Stages, Science Process, Skills and Academic Achievement in Microbiology, *International Journal of Science Education* 24(8), pp. 803-821. DOI: 10.1080/09500690110049150

Ibe, M. and Deutscher, R. 2003. *The Role of the Goldstone Apple Valley Radio Telescope Project in Promoting Scientific Efficacy among Middle and High School Students*. Available at: https://files.eric.ed.gov/fulltext/ED476647.pdf [Accessed: 15 October 2018].

InterAcademies Panel. 2010. Taking IBSE into Secondary Education Report on the conference, York, UK, *Taking IBSE into Secondary Education*, York, 27-29 October. InterAcademies Panel. Available at: http://www.interacademies.net/Publications/25112.aspx [Accessed: 15 October 2014].

Ivankova, N., Creswell, J. and Stick, S. 2006. Using mixed-methods sequential explanatory design: From theory to practice, *Field Methods*, 18(1), pp. 3-20. DOI: 10.1177/1525822X05282260

Jimoyiannis, A. and Komis, V. 2001. Computer simulations in physics teaching and learning: a case study on students' understanding of trajectory motion, *Computers & Education*, 36(2), pp. 183-204. DOI: 10.1016/S0360-1315(00)00059-2

Joffe, H. 2012. Thematic Analysis. In: Harper, D. and Thompson, A. *Qualitative Research Methods in Mental Health and Psychotherapy: A Guide for Students and Practitioners*. Chichester: John Wiley & Sons Ltd.

Johnson, D.W. and Johnson, R.T. 1990. Cooperative Learning and Achievement. In: Sharan, S. Ed. *Cooperative Learning: Theory and Research*. New York, US: Preager Publishers.

Johnson, D.W. and Johnson, R. T. 1999. Making Cooperative Learning Work, *Theory into Practice* 38(2), pp. 67-73. DOI: 10.1080/00405849909543834.

Johnson, R. and Onwuegbuzie, A. 2004. Mixed Methods Research: A Research Paradigm Whose Time Has Come, *Educational Researcher*, 33(7), pp. 14-26. DOI: 10.3102/0013189X033007014

Johnson, B., Onwuegbuzie, A. and Turner, L. 2007. Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), pp. 112-133. DOI: 10.1177/1558689806298224

Joint Council for Qualifications (JCQ) 2010. *Results 2010*. Available at: https://www.jcq.org.uk/examination-results/a-levels [Accessed 14 October 2018].

Joint Council for Qualifications (JCQ) 2012. *Results 2012*. Available at: https://www.jcq.org.uk/examination-results/a-levels [Accessed 14 October 2018].

Joint Council for Qualifications (JCQ) 2014. *Results 2014*. Available at: https://www.jcq.org.uk/examination-results/a-levels [Accessed 14 October 2018].

Joint Council for Qualifications (JCQ). 2015. *Institute of Physics: UK physics A-level entries 2010-15*. Available at: http://www.iop.org/policy/statistics/index.html [Accessed: 26 January 2016]. Joint Council for Qualifications (JCQ) 2016. *Results 2016*. Available at: https://www.jcq.org.uk/examination-results/a-levels [Accessed 14 October 2018].

Joint Council for Qualifications (JCQ) 2018. *Results 2018*. Available at: https://www.jcq.org.uk/examination-results/a-levels [Accessed 14 October 2018].

Jones, S. 1985. Depth Interviewing. In: Walker, R. ed. *Applied Qualitative Research*. Aldershot, UK: Gower, pp. 45-55.

Jones, M.G., Howe, A. and Melissa, R.J. 2000. Gender Differences in Students' Experiences, Interests, and Attitudes toward Science and Scientists, *Science Education* 84(2), pp. 180-192.

Juvonen, J., Espinoza, G. and Knifsend, C. 2012. The Role of Peer Relationships in Student Academic and Extracurricular Engagement, In: Christenson, S., Reschly, A. and Wylie, C. Eds. *Handbook of Research on Student Engagement*. New York, US: Springer, pp. 387-401.

Kagan, S. 1986. Cooperative Learning and Sociological Factors in Schooling. In: Cortes, C. *Beyond Language: Social and Cultural Factors in Schooling Language-Minority Students*. Los Angeles, US: California State University Evaluation, Dissemination and Assessment Center.

Karagiorgi, Y. and Symeou, L. 2005. Translating Constructivism into Instructional Design: Potential and Limitations, *Educational Technology & Society* 8(1), pp. 17-27.

Kelemen-Finan, J., Scheuch, M. and Winter, S. 2018. Contributions from Citizen Science to Science Education: an Examination of a Biodiversity Citizen Science Project with Schools in Central Europe, *International Journal of Science Education* 40(17), pp. 2078-2098. DOI: 10.1080/09500693.152405

Kim. J.S. 2005. The Effects of a Constructivist Teaching Approach on Student Academic Achievement, Self-Concept, and Learning Strategies, *Asia Pacific Education Review* 6(1), pp. 7-19.

Kind, P.M., Jones, K. and Barmby, P. 2007. Developing Attitudes Towards Science Measures, *International Journal of Science Education* 29(7), pp. 871-893. DOI: 10.1080/09500690600909091

Kirschner, P., Sweller, J. and Clark, R. 2006. Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching, *Educational Psychologist*, 41(2), pp. 75-86. DOI: 10.1207/s15326985ep4102 1

Kitzinger, J. 1995. Qualitative Research: Introducing Focus Groups, *Education and Debate* 311. DOI: 10.1136/bmj.311.7000.299

Krapp, A., Hidi, S. and Renninger, K.A. 1992. Interest, Learning and Development. In: Renninger, K.A., Hidi, S., and Krapp, A. Eds. *The Role of Interest in Learning and Development*. Hillsdale US: Lawrence Erlbaum Associates, pp. 3-25.

Lavonen, J., Byman, R., Juuti, K., Meisalo, V. and Uitto A. 2005. Pupil interest in physics: A survey in Finland, *Nordic Studies in Science Education*, 1(2). DOI: 10.5617/nordina.486

Lee, O. and Anderson, C. 1993. Task Engagement and Conceptual Change in Middle School Science Classrooms. *American Educational Research Journal* 30(3), pp. 585-610. DOI: 10.2307/1163304

Leech, N. and Onwuegbuzie, A. 2009. A typology of mixed methods research designs, *Quality and Quantity*, 43(2), pp. 265-275. DOI: 10.1007/s11135-007-9105-3

Levy, P., Lameras, P., McKinney, P. and Ford, N. 2011a. *Pathway UK*. Available at: http://www.pathwayuk.org.uk/what-is-ibse.html. [Accessed: 12 October 2014].

Levy, P., Lameras, P., McKinney, P. and Ford, N. 2011b. *Pathway UK*. Available at: http://www.pathwayuk.org.uk/uploads/9/3/2/1/9321680/_the_features_of_inquiry_learning__theory_r esearch_and_practice_eusubmitted.pdf [Accessed: 20 May 2015].

Lewis, F. and Roche, P. 2009. The Faulkes Telescope Project: Not Just Pretty Pictures. In: Simpson, R.J. and Ward-Thompson, D. *Astronomy: Networked Astronomy and the New Media*. Available at: https://arxiv.org/abs/0902.4809 [15 October 2018].

Li, Q. 2007. Student and teacher views about technology: A tale of two cities?, *Journal of research on Technology in Education*, 39(4), pp. 377-397. DOI: 10.1080/15391523.2007.10782488

Likert, R. 1932. A Technique for the Measurement of Attitudes, Archives of Psychology, 22, pp. 1-55.

Lincoln, Y. and Guba, E. 1985. Naturalistic Inquiry. Beverly Hills, US: Sage Publications.

Linnenbrink, E.A. and Pintrich, P.R. 2003. The Role of Confidence Beliefs in Student Engagement and Learning in the Classroom, *Reading and Writing Quarterly* 19, pp. 119-137. DOI: 10.1080/10573560308223

Loucks-Horsley, S. and Olson, S. 2000. *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning*, Washington DC: National Academies Press.

Lyons, T. 2006. Different Countries, Same Science Classes: Students' Experiences of School Science in their Own Words, *International Journal of Science Education*, 28(6), pp. 591-613. DOI: 10.1080/09500692500339621

Ma, J. and Nickerson, J. 2006. Hands-On, Simulated, and Remote Laboratories: A Comparative Literature Review, *ACM Computing Surveys* 38(3). DOI: 10.1145/1132960.

Mack, L. 2010. *The Philosophical Underpinnings of Educational Research*. Available at: https://secure.apu.ac.jp/rcaps/uploads/fckeditor/publications/polyglossia/Polyglossia_V19_Lindsay.pd f [Accessed: 15 October 2018].

Marsh, S. 2016. *Pressure, No Budgets and 60-hour Weeks: Why Teachers are Striking*. Available at: https://www.theguardian.com/education/2016/jul/05/pressure-no-budgets-and-60-hour-weeks-why-teachers-are-striking [Accessed: 22 October 2018]

Martin, M., Mullis, I., Foy, P. and Stanco, G. 2012. *TIMSS 2011 International Results in Science*. Available at: https://files.eric.ed.gov/fulltext/ED544560.pdf [Accessed: 14 October 2018].

Mäeots, M., Pedaste, M. and Sarapuu, T. 2008. Transforming Student's Inquiry Skills with Computer-Based Simulations', *Eighth IEEE International Conference on Advanced Learning Technologies*, Cantabria, 1-5 July. *IEEE*. Available at: http://ieeexplore.ieee.org [Accessed: 21 March 2015]

McKinnon, D. and Danaia, L. 2008. Remote Telescopes in Education: A Report of an Australian Study. In: *Innovation in Astronomy Education*. Cambridge: Cambridge University Press, pp. 233-242.

McMillan, D. and Chavis, D. 1986. Sense of Community: A Definition and Theory, *Journal of Community Psychology* 14(1), pp. 6-23.

Merriam, S. 1988. Case Study Research in Education. San Francisco, US: Jossey-Bass Publishers.

Minner, D., Levy, A. and Century, J. 2010. Inquiry-based Science Instruction – what is it and does it matter? Results from a Research Synthesis years 1984 to 2002, *Journal of Research in Science Teaching*, 47(4), pp. 474-496.

Morgan, D. and Spanish, M. 1984. Focus Groups: A New Tool for Qualitative Research, *Qualitative Sociology*, 7(3), pp. 253-270.

Muijs, D. 2004. Doing Quantitative Research in Education with SPSS. London: SAGE Publications.

Mujtaba, T. and Reiss, M. 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), pp. 1824-1845. DOI: 10.1080/09500693.2012.762699.

Mulhall, P and Gunstone, R. 2012. Views about Learning Physics Held by Physics Teachers with Differing Approaches to Teaching Physics, *Journal of Science Teacher Education* 23(5), pp. 429-449. DOI: 10.1007/s10972-012-9291-2

Murray, I. and Reiss, M. 2005. The student review of the science curriculum', *School Science Review*, 87(318). Available at: http://eprints.ioe.ac.uk/449/ [Accessed: 12 May 2015].

NAHT Accountability Commission. 2018. *Improving School Accountability*. Available at: https://www.naht.org.uk/news-and-opinion/news/structures-inspection-and-accountability-news/improving-school-accountability/ [Accessed: 10 December 2018]

Nakhleh, M., Auberry, D. and Eichinger, D. 2000. Evaluating Computer Lab Modules for Large Biology Courses. *Journal of Computers in Mathematics and Science Teaching* 19(3), pp. 253-275.

National Assessment of Educational Progress (NAEP). 2018. *Survey Questionnaires*. Available at: https://nces.ed.gov/nationsreportcard/experience/survey_questionnaires.aspx#student [Accessed: 11 December 2018]

National Audit Office. 2016. *Training new Teachers*. Available at: https://www.nao.org.uk/wp-content/uploads/2016/02/Training-new-teachers.pdf [Accessed: 22 October 2018]

National Centre for Educational Statistics (NCES). 2012. *Questionnaires for Students, Teachers and School Administrators*, Available at: https://nces.ed.gov/nationsreportcard/bgquest.aspx#student [Accessed: 28 January 2016].

National Research Council. 1996. *National Science Education Standards*. Washington DC: National Academy Press, p. 23.

National Research Council. 2001. *Astronomy and Astrophysics in the New Millennium*. Washington DC, US: The National Academy Press. DOI: 10.17226/9839

Nokelainen, P. 2006. An Empirical Assessment of Pedagogical Usability Criteria for Digital Learning Material with Elementary School Students. *Journal of Educational Technology and Society* 9(2), pp. 178-197.

Nowell, L., Norris, J. White, D. and Moules, N. 2017. Thematic Analysis: Striving to Meet the Trustworthiness Criteria, *International Journal of Qualitative Methods* 16(1-13). DOI: 10.1177/1609406917733847

OECD. 2014. *Improving Schools in Wales: An OECD Perspective*. Available at: http://www.oecd.org/education/Improving-schools-in-Wales.pdf [Accessed: 09 December 2018]

Ofsted. 2011. *Successful Science: An Evaluation of Science Education in England 2007-2010.* Available at:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/413 805/Summary_-_Successful_science.pdf [Accessed: 15 October 2018].

Ofsted. 2013. *Maintaining curiosity: a survey into science education in schools*. Available at: https://www.gov.uk/government/publications/maintaining-curiosity-a-survey-into-science-education-in-schools [Accessed: 26 January 2016].

Organisation for Economic Co-operation and Development (OECD). 2006. *Global Science Reform: Evolution of Student Interest in Science and Technology Studies Policy Report*. Available at: http://www.oecd.org/science/sci-tech/36645825.pdf [Accessed: 22 October 2018]

Osborne, J. and Collins, S. 2001. Pupils' views of the role and value of the science curriculum: A focus-group study, *International Journal of Science Education*, 23(5), pp. 441-67. DOI: 10.1080/09500690010006518

Osborne, J. and Dillon, J. 2008. *Science education in Europe: Critical reflections*. London: The Nuffield Foundation. Vol. 13.

Osborne, J., Simon, S. and Collins, S. 2003. Attitudes towards science: a review of the literature and its implications, *International Journal of Science Education*, 25, pp. 1049-79.

Osborne, J., Simon, S., and Tyler, R. 2009. Attitudes towards school science: An update, *Annual Meeting of the American Educational Research Association*, California, 13-17 April. American Educational Research Association. Available at: http://science.teachingjournalclub.org/wp-content/uploads/2012/06/Attitudes-Toward-Science-OsborneSimon-Tytler.pdf [Accessed: 26 April 2015].

Osterman K. 2000. Students' Need for Belonging in the School Community. *Review of Educational Research* 70(3), pp. 323-367. DOI: 10.3102/00346543070003323

Palmer, D. 2009. Student Interest Generated During an Inquiry Skills Lesson, *Research on Motivation and Student Interest in Science*, 46(2), pp. 147-165. DOI: 10.1002/tea.20263.

Panitz, T. 2002. The Motivational Benefits of Cooperative Learning, *New Directions for Teaching and Learning* 78, pp. 59-67. DOI: 10.1002/tl.7806

Parkinson, J., Hendley, D., Tanner, H., Stables, A. 1998. Pupils' Attitudes to Science in Key Stage 3 of the National Curriculum: A Study of Pupils in South Wales, *Research in Science and Technological Education*, 16(2), pp. 165-176. DOI: 10.1080/0263514980160206

Parlett, M.B. and Hamilton, D. 1972. *Evaluation as Illumination: A New Approach to the Study of Innovatory Programs*. Available at: https://files.eric.ed.gov/fulltext/ED167634.pdf [Accessed: 11 October 2018]

Pavlov, I.P. 1902. The Work of the Digestive Glands. London: Griffin.

Pedaste, M., Mäoets, M., Siiman, L., de Jong, T., Riesen, S., Kamp, E., Manoli, C., Zacharia, Z. and Tsourlidaki, E. 2015. Phases of inquiry-based learning: Definitions and the inquiry cycle, *Educational Research Review*, 14, pp. 47-61. DOI: 10.1016/j.edurev.2015.02.003

Pell, T. and Jarvis, T. 2001. Developing attitude to science scales for use with children of ages from five to eleven years. *International Journal of Science Education* 23(8), pp. 847-862. (check ref)

Peters, J., Stylianidou, F., Ingram, C., Malek, R., Reiss, M. and Chapman, S. 2006. *Evaluation of Einstein Year*. Available at:

https://s3.amazonaws.com/academia.edu.documents/30720849/file_39565.pdf?AWSAccessKeyId=A KIAIWOWYYGZ2Y53UL3A&Expires=1538833937&Signature=kfOSLvQ9XKeyQWuM%2BvEm AqycJsk%3D&response-content-

disposition=inline%3B%20filename%3DEvaluation_of_Einstein_Year.pdf [Accessed: 06 October 2018].

Piaget, J. 1977. *The development of thought: Equilibration of cognitive structures*. Oxford, England: Viking.

Pianta, R., Hamre, B. and Allen, J. 2012. Teacher-Student Relationships and Engagement:Conceptualizing, Measuring, and Improving the Capacity of Classroom Interactions. In: Christenson,S., Reschly, A. and Wylie, C. (Eds). *Handbook of Research on Student Engagement*. New York:Springer, pp. 365-386

Presser, S., Couper, M., Lessler, J., Martin, E, Martin, J., Rothgeb, J. and Singer, E. 2004. Methods for Testing and Evaluating Survey Questions, *Public Opinion Quarterly*, 68(1), pp. 109-130. DOI: 10.1093/poq/nfh008

Prince, M. and Felder, R. 2006. Inductive Teaching and Learning Methods: Definitions, Comparisons, and Research Bases, *Journal of Engineering Education*, 95(2), pp. 123-138.

Protopsaltis, A. and Specht, M. (Eds.) 2013. *Introduction on Inquiry-Based Learning (IBL)*. Available at: https://www.scribd.com/document/333143157/into-IBL [Accessed: 10 December]

Punch, K. 2009. Introduction to Research Methods in Education, London: SAGE Publications.

Quinn, C. 1996. Pragmatic Evaluation: Lessons from Usability. *The 13th Annual Conference of the Australasian Society for Computers in Learning and Tertiary Education*. Available at: http://www.ascilite.org/conferences/adelaide96/papers/18.html [Accessed: 09 August 2018].

Reeves, T. 1994. Evaluating What Really Matters in Computer-Based Education. In: Wild, M. and Kirkpatrick, D. eds. *Computer Education: New Perspective*. Perth: Mathematics, Science & Technology Education Centre, Edith Cowan University, pp. 219-246.

Reid, N. 2006. Thoughts on Attitude Measurement, *Research in Science and Technological Education* 24(1), pp. 3-27. DOI: 10.1080/02635140500485332

Reid N. and Skryabina, E. 2002. Attitudes towards Physics, *Research in Science and Technological Education* 20(1), pp. 67-81. DOI: 10.1080/02635140220130939.

Reige, A. 2003. Validity and Reliability Tests in Case Study Research: A Literature Review with 'Hands-on' Applications for each Research Phase, *Qualitative Market Research* 6(2), pp. 75-86. DOI: 10.1108/13522750310470055

Reigeluth, C. 1989. Educational Technology at the Crossroads: New Mindsets and New Directions, *Educational Technology Research and Development*, 37(1), pp. 67-80. DOI: 10.1007/BF02299047

Reigeluth, C. 1999. What is Instructional-Design Theory and How is it Changing? In: Reigeluth, C. *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory, Volume 2.* New Jersey: Lawrence Erlbaum Associates.

Richwine, P. 2007. *The Impact of Authentic Science Inquiry Experiences Studying Variable Stars on High School Students' Knowledge and Attitudes about Science and Astronomy and Beliefs Regarding the Nature of Science*. PhD Thesis, University of Arizona.

Rijsdijk, C. 2000. Using Astronomy as a Vehicle for Science Education, *Publications of the Astronomical Society of Australia* 17(2), pp. 156-161. DOI: 10.1071/AS00156

Ritchie, J. 2003. The Applications of Qualitative Methods to Social Research. In: Ritchie, J. and Lewis, L eds. *Qualitative research practice: A guide for social science students and researchers*. London: SAGE Publications, pp. 24-46.

Rittmayer, A. and Beier, M. 2009. Confidence in STEM. In: Bogue, B. and Cady, E. Eds. Applying *Research to Practice (ARP) Resources*. Available at:

https://www.engr.psu.edu/awe/secured/director/assessment/Literature_Overview/PDF_overviews/AR P SelfEfficacy Overview 122208.pdf [Accessed: 15 October 2018].

Roblyer, M., Edwards, J. and Havriluk, M. 1997. *Integrating educational technology into teaching* (2nd ed.). Upper Saddle River, USA: Prentice-Hall.

Robson, C. 1993. *Real world research: A resource for social scientists and practitioner-researchers.* Oxford: Blackwell Publishers.

Rocard, M., Csermely, P., Jorde, D., Lenzen, D., Walberg-Henriksson, H. and Hemmo, V. 2007. *Rocard Report: Science Education Now: a Renewed Pedagogy for the Future of Europe, European Commission Research and Innovation e-Library*. Available at: http://ec.europa.eu/research/science-society/document_library/pdf_06/report-rocard-on-science-education_en.pdf [Accessed: 24 April 2015].

Rodd, M., Reiss, M. and Mujtaba, T. 2014. Qualified, But not Choosing STEM at University: Unconscious Influences on Choice of Study, *Canadian Journal of Science, Mathematics and Technology Education* 14(4), pp. 330-345. DOI: 10.1080/14926156.2014.938838.

Roeser, R., Midgley, C. and Urdan, T. 1996. Environment and Early Adolescents' Psychological and Behavioural Functioning in Case: The Mediating Role of Goals and Belonging. *Journal of Educational Psychology* 88(3), pp. 408-422. DOI: 10.1037/0022-0663.88.3.408

Roorda, D., Koomen, H., Spilt, J. and Oort, F. 2011. Influence of Affective Teacher-Student Relationships on Students' School Engagement and Achievement: A Meta-Analytic Approach, *Review of Educational Research* 81(2), pp. 493-529.

Rutten, N., Joolingen, W. and van der Veen, J. 2012. The Learning Effects of Computer Simulations in Science Education. *Computers & Education* 58(1), pp. 136-153. DOI: 10.1016/j.compedu.2011.07.017

Ryan, A. 2006. Post-Positivist Approaches to Research. In: *Researching and Writing your Thesis: A Guide for Postgraduate Students*, Maynooth Adult and Community Education, pp. 12-26. Available at: http://eprints.maynoothuniversity.ie/874/1/post-positivist_approaches_to_research.pdf [Accessed: 15 October 2018]

Sadeh, I. and Zion, M. 2012. Which Type of Inquiry Project do High School Biology Students Prefer: Open or Guided?, *Research in Science Education* 42(5), pp. 831-848. DOI: 10.1007/s11165-011-9222-9

Sadler, P., Gould, R., Leiker, S., Antonucci, P. and Kimberk, R. 2001. MicroObservatory Net: A Network of Automates Remote Telescopes Dedicated to Educational Use, *Journal of Science Education and Technology* 10(1) pp. 39-55. DOI: 10.1023/A:1016668526933

Salimpour, S., Bartlett, S., Fitzgerald, M., Cutts, R.K., James, R., Miller, S., Danaia, L., Cabezon, S., Faye, M., Baudouin, C., Birkenbauma, D., Anjos, S., Wu, Q., Chu, H. and Slater, E. 2018. Robotic Telescopes and Student Research in the School Curriculum around the OECD Countries. In: Fitzgerald, M., James, M., Buxner, C.R. and White, S. Eds. *Robotic Telescopes, Student Research and Education (RTSRE)*. San Diego, US, 18-21 June, 2017, pp. 33-40. DOI: 10.32374/rtsre.2017.003

Sandberg, J. 2005. How Do We Justify Knowledge Produced Within Interpretive Approaches? *Organizational Research Methods* 8(1). Available at: https://www.researchgate.net/publication/43448121_How_Do_We_Justify_Knowledge_Produced_W ithin Interpretive Approaches [Accessed: 21 December 2016].

Sarabando, C., Cravino, J. and Soares, A. 2014. Contribution of a Computer Simulations to Students' Learning of the Physics Concepts of Weight and Mass. *Procedia Technology* 13. pp. 112-121: DOI: 10.1016/j.protcy.2014.02.015

Sauter, M., Uttal, D., Rapp, D., Downing, M. and Jona, K. 2013. Getting real: the authenticity of remote labs and simulations for science learning, *Distance Education*, 34(1), pp. 37-47. DOI: 10.1080/01587979.2013.770431

Schau, C., Stevens, J., Dauphinee, T.L. and Del Vecchio, A. 1995. The Development and Validation of the Survey of Attitudes towards Statistics, *Educational and Psychological Measurement* 55(5), pp. 868-875. DOI: 10.1177/0013164495055005022.

Schiefele, U. 1991. Interest, Learning and Motivation, *Educational Psychologist* 26(3-4), pp. 299-323. DOI: 10.1080/00461520.1991.9653136

Schwab, J. 1960. What do Scientists do?, Behavioural Science, 5(1), pp. 1-27.

Schwab, J. 1962. The teaching of science as enquiry, in Schwab, J. and Brandwein, P. F. (eds.) The Teaching of Science. Cambridge, U.S.: Harvard University Press, pp. 3-103.

Scriven, M. 1991. Pose and Cons about Goal-Free Evaluation. *Evaluation Practice* [Online] 12(1), pp. 55-63. Available at: http://www.sciencedirect.com/science/article/pii/088616339190024R [Accessed: 03 January 2017].

Simons, K. D. and Ertmer, P. A. 2005. Scaffolding disciplined inquiry in problem-based environments, *International Journal of Learning*, 12(6), pp. 297-305. DOI: 10.18848/1447-9494/CGP/v12i06/47900

Siorenta, A. and Jimoyiannis, A. 2008. Physics Instruction in Secondary Schools: An Investigation of Teachers' Beliefs towards Physics Laboratory and ICT, *Research in Science and Technological Education* 26(2), pp. 185-202. DOI: 10.1080/02635140802037328

Sjøberg, S. 2000. Interesting all children in 'science for all'. In Millar, R., Leach, J and Osborne, J. Eds. *Improving Science Education*. The Contribution of Research. Buckingham: Open University Press.

Sjøberg, S. and Schreiner, C. 2010. *The ROSE project: An overview and key findings*. Oslo: University of Oslo.

Slater, S., Slater, T., Heyer, I. and Bailey, J. 2015. *Conducting Astronomy Education Research: An Astronomer's Guide*. 2nd Edition. Wyoming, U.S.: Pono Publishing.

Smetana, L. and Bell, R. 2012. Computer Simulations to Support Science Instruction and Learning: A critical review of the literature. *International Journal of Science Education* 34(9), pp. 1337-1370. DOI: 10.1080/09500693.2011.605182

Songer, N. B., Lee, H. S. and McDonald, S. 2003. Research towards and expanded understanding of inquiry science beyond one idealized standard, *Science Education*, 87(4), pp. 490-516.

Spall, K., Stanisstreet, M., Dickson, D. and Boyes, E. 2004. Development of School Students' Constructions of Biology and Physics, *International Journal of Science Education* 26(7), pp. 787-803. DOI: 10.1080/0950069032000097442.

Spector, P. 1992. *Summated Rating Scale Construction: An Introduction*. California, US: Sage Publications.

Spronken-Smith, R. and Walker, R. 2010. Can inquiry-based learning strengthen the links between teaching and disciplinary research?, *Studies in Higher Education*, 35(6), pp. 723-740. DOI: 10.1080/03075070903315502

Squires, D. and Preece, J. 1999. Predicting Quality in Educational Software: Evaluating for Learning, Usability and the Synergy between them. *Interacting with Computers* 11(5), pp. 467-483. DOI: 10.1016/S0953-5438(98)00063-0.

Stake, R.E. 1978. The Case Study Method in Social Inquiry, *Educational Researcher* 7(2). DOI: 10.3102/0013189X007002005

Stake, R.E. 1995. The Art of Case Study Research. Thousand Oaks: SAGE Publications

Stake, R.E. 2010. Qualitative Research: Studying How Things Work. New York, US: Guildford Press.

Stipek, D.J. 2002. *Motivation to Learn: Integrating Theory and Practice*, 4th Ed. Boston, US: Allyn-Bacon.

Strauss, A. and Corbin, J. 1998. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, US: SAGE Publications.

Straw, S., MacLeod, S. and Hart, R. 2012. *Evaluation of the Wellcome Trust Camden STEM Initiative*. Slough: NFER.

Stuckey, M., Hofstein, A., Mamlok-Naaman, R. and Eilks, I. 2013. The Meaning of 'Relevance' in Science Education and its Implications for the Science Curriculum, *Studies in Science Education* 49(1), pp. 1-34. DOI: 10.1080/03057267.2013.802463

Symon, G. and Cassell, C. 2012. *Qualitative Organizational Research: Core Methods and Current Challenges*. London: SAGE Publications.

Taasoobshirazi, G., Zuiker, S., Anderson, K. and Hickey, D. 2006. Enhancing Inquiry, Understanding, and Achievement in an Astronomy Multimedia Learning Environment, *Journal of Science Education and Technology* 15(5), pp. 383-395. DOI: 10.1007/s10956-006-9028-0

Tafoya, E., Sunal, D. and Knecht, P. 1980. Assessing Inquiry Potential: A Tool for Curriculum Decision Makers, *School Sciences and Mathematics*, 80(1) pp. 43-48.

Teddlie, C. and Tashakkori, C. 2009. Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioural Sciences. 7th Edition. Thousand Oaks, US: SAGE Publications.

Tellis, W. 1997. Application of a Case Study Methodology. The Qualitative Report 3(3), pp. 1-19.

Temitope, B. and Christy, O. 2015. Influence of Peer Group on Academic Performance of Secondary School Students in Ekiti State, *International Journal of Innovative Research and Development* 4(1), pp. 324-331.

Thomas, D.R. 2006. A General Inductive Approach for Analyzing Qualitative Evaluation Data, *American Journal of Evaluation* 27(2). DOI: 10.1177/1098214005283748

Thomson Committee Report. 1918. *The Position of Natural Science in the Education System of Great Britain: Report of the Committee of Enquiry*. London: HMSO.

Tomlinson, C. 1995. Differentiation Instruction for Advanced Learners in the Mixed-Ability Middle School. *ERIC Digest E536*. Available at: https://files.eric.ed.gov/fulltext/ED389141.pdf [Accessed: 28 August 2018].

Townsend, A. 2015. *Educative Curricula and Improving the Science PCK of Teachers in Middle School Settings in Rural and Remote Australia*. PhD Thesis, Edith Cowan University.

Treisman, U. 1992. Studying students studying calculus: a look at the lives of minority mathematics students in college', *The College Mathematics Journal*, 23(5), pp. 362-372.

Tripney, J., Bangpan, M. and Newman, M. 2010. *Factors Influencing Young People (aged 14-19) in Education Subject Choices: A Systematic Review of the UK Literature*. Available at: https://www.researchgate.net/publication/265086031_Factors_Influencing_Young_People_aged_14-19_in_Education_about_STEM_Subject_Choices_A_systematic_review_of_the_UK_literature [Accessed: 14 October 2018]

Trnova, E. and Trna, J. 2015. Motivational Effectiveness of a Scenario in IBSE, *Procedia – Social and Behavioural Sciences*, 167, pp. 184-189. DOI: 10.1016/j.sbspro.2014.12.660

Trochim, W. 2006. *The Research Methods Knowledge Base*, 2nd Edition. *Social Research Methods*. Available at: http://www.socialresearchmethods.net/ [Accessed: 15 January 2016].

Trumper, R. 2006a. Factors Affecting Junior High School Students' Interest in Physics, *Journal of Science Education and Technology*, 15(1), pp. 47-58.

Trumper, R. 2006b. Teaching Future Teachers Basic Astronomy Concepts – Seasonal Changes – at a Time of Reform in Science Education, *Journal of Research in Science Teaching*, 43(9), pp. 879-906.

Tsai, C. 2002. Nested Epistemologies: Science Teachers' Beliefs of Teaching, Learning and Science, *International Journal of Science Education* 24(8), pp. 771-783. DOI: 10.1080/09500690110049132

Tuan, H., Chin, C., Tsai, C. and Cheng, S. 2005. Investigating the effectiveness of inquiry instruction on the motivation of different learning styles students, *International Journal of Science and Mathematics Education*, 3(4), pp. 541-566.

Turner, P.R., Fowler, K., Wick, D., Ramsdell, M., Gotham, G., Glasgow, E. and French, C. 2007. BOCES-University Partnership as a Model for Educational Outreach: K-16 STEM Professional Development. *Math and Science Symposium*. Knoxville, US, October. Available at: https://www.clarkson.edu/sites/default/files/2018-10/Peter-BOCESpaper.pdf [Accessed: 23 November 2018]

Tymms, P., Bolden, D.S. and Merrell, C. 2008. Science in English primary schools: trends in attainment, attitudes and approaches', *Perspectives on Education: Primary Science*, 1, pp. 19-42.

UCAS. 2018. 2018 Cycle Applicant Figures – June Deadline. Available at: https://www.ucas.com/corporate/data-and-analysis/ucas-undergraduate-releases/2018-cycle-applicant-figures-june-deadline [Accessed: 17 October 2018].

UK Government. 2017. Industrial Strategy: Science and STEM Skills: Government Response to the Committee's Thirteenth Report of Session 2016-17. Available at: https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/335/33502.htm [Accessed: 11 December 2018]

Urdan, T. and Schoenfelder, E. 2006. Classroom Effects on Student Motivation: Goal Structures, Social Relationships and Competence Beliefs, *Journal of School Psychology* 44(5), pp. 331-349. DOI: 10.1016/j.jsp.2006.04.003

Vaughn, S., Schumm, J.S. and Sinagub, J. 1996. *Focus Group Interviews in Education and Psychology*. Thousand Oaks: SAGE Publications.

Verma, G. and Mallick, K. 1999. Researching Education. London: Farmer Press.

Vennix, J., den Brok, P. and Taconis, R. 2018. Do Outreach Activities in Secondary STEM Education Motivate Students and Improve their Attitudes towards STEM?, *International Journal of Science Education* 40(11), pp. 1263-1283. DOI: 10.1080/09500693.2018.1473659

Vygotsky, L. S. 1978. *Mind in Society: The development of higher Psychological Processes*. Cambridge, Massachusetts: Harvard University Press.

Ward, H. 2018. *Secondary Applicants Down and Four Other New Teacher Training Stats*. Available at: https://www.tes.com/news/secondary-applicants-down-and-four-other-new-teacher-training-stats [Accessed: 22 October 2018]

Webb, M. 2010. Technology-mediated Learning. In: Osborne, J. and Dillon, J. eds. *Good Practice in Science Teaching: What research has to say.* 2nd ed. Maidenhead: Open University Press, pp. 158-182

Wellcome Trust. 2011. *Exploring Young People's Views on Science Education*. Available at: https://wellcome.ac.uk/sites/default/files/wtvm052732_0.pdf [Accessed: 14 October 2018].

Wellcome Trust. 2015. *Review of the Extent to which Ofsted Reports Mention Science*. Available at: https://wellcome.ac.uk/sites/default/files/extent-to-which-ofsted-reports-mention-science-wellcome-sep15.pdf [Accessed: 14 October 2018].

Wellcome Trust. 2016a. *Wellcome Trust: SET Development*. Available at: https://wellcome.ac.uk/sites/default/files/science-education-tracker-wellcome-feb16.pdf [Accessed: 09 October 2018].

Wellcome Trust. 2016b. *Executive Summary – NFER Teacher Voice Omnibus Survey Data*. Available at: https://wellcome.ac.uk/sites/default/files/nfer-teacher-voice-omnibus-survey-data-apr16.pdf [Accessed: 15 October 2018].

Welsh Government. 2015. *The National Science Academy STEM Enrichment Strategic Plan 2015-18*. Available at: https://gov.wales/docs/det/publications/150701-nsa-strategy-en.pdf [Accessed 11 December 2018]

Welsh Government. 2017a. School Performance Measurement: Contribution Values of Qualifications. Available at: https://beta.gov.wales/school-performance-measurement-contribution-values-qualifications [Accessed: 07 December 2018]

Welsh Government. 2017b. *My Local School: Sharing Information about Local Schools*. Available at: http://mylocalschool.gov.wales/?lang=en [Accessed: 11 October 2018].

Welsh Government: Department for Education and Skills. 2008. *Science in the national curriculum for Wales*. Available at: http://learning.gov.wales/docs/learningwales/publications/140624-science-in-the-national-curriculum-en.pdf [Accessed: 26 January 2016].

Wieman, C. 2007. Why Not Try a Scientific Approach to Science Education?, *Change: The Magazine of Higher Learning*, 39(5), pp. 9-15.

Wightwick, A. 2018. *Schools in Wales will no longer be Judged on How Many Pupils get Five GCSEs Grades A* to C.* Available at: https://www.walesonline.co.uk/news/education/schools-wales-no-longer-judged-14675376 [Accessed: 07 December 2018]

Williams, C., Stanisstreet, M., Spall, K., Boyes, E. and Dickson, D. 2003. Why aren't secondary students interested in physics? *Physics Education*, 38(4) p. 324. DOI: 10.1088/0031-9120/38/4/306

Wolcott, H. 2002. Writing up Qualitative Research...Better, *Qualitative Health Research* 12(1). DOI: 10.1177/1049732302012001007

Wolf, S. and Fraser, B. 2008. Learning Environment, Attitudes and Achievement among Middle-School Science Students Using Inquiry-Based Laboratory Activities, *Research in Science Education* 38(3), pp. 321-341.

Wood, D., Bruner, J. and Ross, G. 1976. The Role of Tutoring in Problem Solving, *Journal of Child Psychology and Psychiatry*, 17(2), pp. 89-100. DOI: 10.1111/j.1469-7610.1976.tb00381

Wragg, E. C. 2012. *An Introduction to Classroom Observation*, 2nd ed. Abingdon: Routledge. Available at: https://www.dawsonera.com/ [Accessed: 6 April 2016].

Wynarczyk, P. and Hale, S. 2009. *Improving Take up of Science and Technology Subjects in Schools and Colleges: A Synthesis Review*. Available at: https://esrc.ukri.org/files/public-engagement/public-dialogues/full-report-improving-the-take-up-of-science-and-technology/ [Accessed: 14 October 2018].

Yager, R.E. 1991. The Constructivist Learning Model: Towards Real Reform in Science Education, *Science Teacher* 58(6), pp. 52-57.

Yazan, B. 2015. Three Approaches to Case Study Methods in Education: Yin, Merriam and Stake. *The Qualitative Report* 20(2), pp. 134-152.

Yilmaz, K. 2013. Comparison of Quantitative and Qualitative Research Traditions: Epistemological, Theoretical and Methodological Differences, *European Journal of Education* 48(2), pp. 311-325. DOI: 10.1111/ejed.12014

Yin, R. 2009. *Case Study Research: Design and Methods*. 4th Ed. Thousand Oaks: SAGE Publications. Applied Social Research Methods Series.

Yin, R. 2014. *Case Study Research Design and Methods*. 5th Edition. Thousand Oaks, US: SAGE Publications.

Zainal, Z. 2007. Case study as a research method. *Jurnal Kemanusiaan* 9, pp. 1-6. Available at: https://jurnalkemanusiaan.utm.my/index.php/kemanusiaan/article/view/165/158 [Accessed: 23 June 2016].

Zeilik, M., Schau, C., Mattern, N., Hall, S., Teague, K.W. and Bisard, W. 1997. Conceptual Astronomy: A Novel Model for Teaching Postsecondary Science Courses, *American Journal of Physics* 65, pp. 987-996. DOI: 10.119/1.18703

Zeilik, M., Schau, C. and Mattern, N. 1999. Conceptual Astronomy II: Replicating Conceptual Gains, Probing Attitude Changes across Three Semesters, *American Journal of Physics* 67, pp. 923-927. DOI: 10.119/1.19151

Zion, M. and Mendelovici, R. 2012. Moving from Structured to Open Inquiry: Challenges and Limits, Science Education International 23(4), pp. 383-399.

Appendices

Appendix 1 – Down2Earth 'Recreate a Crater' Resource

Recreate a Crater: Student Worksheet

Research Scenario

In this activity, you will use the Down2Earth Impact Calculator to recreate particular impact craters found in the Solar System. You will explore how changes in the properties of an impactor affects the size of the crater that is produced upon impact and whether it is possible to create impact craters on Earth that resemble some of those found on other Solar System bodies.

Activity

Equipment

- Computer with internet access
- Calculator
- Pen

Instructions

Planning your Investigation

Step 1: Launch the Down2Earth Impact Calculator.

Step 2: In this activity, you will be provided with the **projectile density** and **diameter** and the **target density** of Earth. Look at the input parameters of the Down2Earth Impact Calculator and consider what other parameters must be considered in your investigation. Use the space below to make a note of what parameters you expect will need to be recorded throughout your investigation.

Step 3: Now that you have considered the parameters involved in this activity, construct a **hypothesis**, below, about what you expect to find in your investigation, i.e. how changing certain parameters affects the resulting impact crater.

Recreating the Impact Craters

We can now begin the investigation and attempt to recreate the five impact craters you have been introduced to.

Step 4: Begin by attempting to construct two impactors that produce craters close to the sizes of the Aorounga and Barringer impact craters on Earth.

We will assume the original body was a comet, composed of **ice** and a target material of **sedimentary rock.** Apply this information, and what is already provided in Table 1, to complete the rest of the table.

Table 1 – Impact Craters on Earth						
Actual Crater	Simulated Crater Width (m)	Projectile Diameter (m)	Trajectory Angle (degrees)	Velocity (km/s)		
Aorounga (12,000 m)		2000				
Barringer (1,186 m)		200				

When you are happy with the parameters of your impactor, use the drop down menu on the **Crater Size** tab of the Impact Calculator to locate the Aorounga and Barringer craters on Google Maps.

Step 5: Place the craters of your impactors next to their corresponding craters and print-screen the image of the two, side-by-side, in order to see the similarity.

You will notice that the Barringer Meteor crater appears to be square shaped. This is due to the geology of the target material and the relatively small impactor that caused it. If the impactor had been larger, it would have obliterated these shape controls.

Now you have successfully recreated two of Earth's famous impact craters, let's try and bring some of the craters found elsewhere in the Solar System, to Earth.

Step 6: This time we will again assume a target material of **sedimentary rock** but change the projectile density to **iron**.

Using this information and again what you are provided with in Table 2, construct your own impactors that will result in craters of widths corresponding to the three craters found on the Moon, Mars and Venus. Record your results in Table 2 below.

Table 2 – Impact Craters on the Moon, Mars and Venus						
Actual Crater	Actual CraterSimulated CraterProjectileTrajecWidth (m)Diameter (m)Angle (d)					
Rabe (108,000 m)		7000				
Danilova		3000				

(50,000 m)		
Pythagoras (120,000 m)	9000	

Step 7: Now experiment with creating these craters using a projectile density of **ice**. Make a note of what you find in the space below.

Now that you have explored the parameters involved in producing impact craters, you can answer the questions, below, that apply some of the scientific concepts and mathematical formulae from your curriculum to this particular scenario.

Questions

Now that you have experimented with the parameters of impacts and obtained all the essential facts and figures, you can apply them to the research scenario.

The following questions consider how we can calculate some of the energies involved in impact events and uncover further information about the properties of impactors.

Data Interpretation

1/. a). When impactors strike the surface of a planetary body, they will exert a certain amount of force. What does Newton's third law tell you about how the surface of the planetary body will react to this impact?

b). What kind of force describes the one that is exerted by the planetary body?

2/. a). Assume an impactor to have a velocity of 30.0 km s⁻¹ and a mass of 9.05 x 10^{11} kg. What is the momentum of this object?

b). What is the kinetic energy of this object?

c). The impact energy of this object was equal to 3.98 x 10²⁰ J. How much energy has been lost?

d). What form do you think this energy was lost in?

e). This same impactor also has a diameter of 1,200 m. Assuming the object to be spherical, calculate its density.

f). What material does this value of density correspond to?

3/. From your findings in Step 7 of the activity, what can you conclude about how the density of an impactor influences the size of the crater it produces?

4/. What properties possessed by a planet or moon do you think will affect the size of the crater that is produced in an impact event?

Now that you have completed your investigation and interpreted your findings, you can refer back to the initial research question and your hypothesis.

The following questions reflect on the process of your investigation and what you can conclude from it.

5/. Based on your findings, are you able to conclude how the width of the crater is affected by changing the following (assume all other parameters remain unchanged):

a). Increasing the trajectory angle

b). Reducing the projectile velocity

c). Increasing the projectile diameter

d). Increasing the projectile density

6/. Do your conclusions match your original hypothesis at the beginning of the investigation?

Appendix 2 – Hubble's Expansion of the Universe Resource Extract

The Expansion and Age of the Universe



Using type Ia supernovae to calculate the expansion and age of our Universe

We can use type Ia supernovae to help us determine how quickly our Universe is expanding and also, how old it is! This activity involves combining you own data with other scientific data to calculate these values.

If you have completed the Spotting a Supernova activity, you can use your results to **contribute to the data** in this activity.

If you haven't, you can still complete this activity with the data provided.

Measuring Peak Magnitudes of Type Ia Supernovae

Now that you are familiar with carrying out photometry, let's take a look at some data from some other type la supernova targets.

Step 1: First of all, you will need to open the Excel Spreadsheet with a table prepared for you to record your results. Here you will find the following columns:

- Name of Supernova
- Apparent Magnitude
- Absolute Magnitude
- Distance (2 columns)
- Host Galaxy
- Radial Velocity

Step 2: Go to the Excel spreadsheet '**Type Ia Supernovae**', in here you should find 15 sheets of data. Each of these sheets contain data for a particular supernova target and are named accordingly.

Step 3: On the first sheet (1994S), you will see a list of data and a light curve that displays Magnitude on the y-axis and MJD on the x-axis.

Note: This magnitude value corresponds to apparent magnitude.

MJD is an acronym for **Modified Julian Date** and is a common dating system that is used by astronomers. For an explanation of MJD, see Appendix 2.

Step 4: Remember, magnitude scales are reversed, so brighter objects have lower magnitude values

Figure 1 – Identifying peak magnitude of a supernova light curve

	A	В	С	D	E	F	G	Н	1.1	J	K	1
1	MJD	Magnitude	Error	Filter								
2	49513.7	15.04	0.05	v								
3	49514.8	14.96	0.03	V								5
4	49515.7	14.93	0.03	V		Su	perno	va Ligh	t Curv	e		
5	49517.7	14.84	> 0.06	V								
6	49519.7	14.85	0.04	V				MJD				
7	49529.7	15.21	0.04	V	49510	495	20 4	19530	49540	49550	49560	
8	49536.6	15.65	0.06	V	14.8	<u> </u>						
9	49538.7	15.76	0.05	V	15	**						
10	49541.7	15.89	0.03	V	15.2			+				
11	49545.7	16.08	0.03	v	15.4							
12	49547.7	16.19	0.06	v .	15.8			•	+			
13	49553.7	16.65	0.1	v	ž 16				- T 4			
14					16.2							
15					16.6					•		
16					16.8							
17												_

Step 5: Now you will need to **identify the peak magnitude**. This is the point where the supernova is at its **brightest**. You can do this either by looking at your graph or by reading from the listed data. When you have done this, add the value to the corresponding column in your table of results along with the name of this supernova. An example is illustrated in Figure 1.

Note: Remember, the peak magnitude will be the lowest value.

Step 6: Repeat steps 3 to 5 for each of the supernova targets on each sheet in the spreadsheet and complete the **Magnitude column** in your table.

Calculating Distance with Supernovae

You should have now completed the columns under **Supernova Name** and **Magnitude**. Next you will need to obtain the values for **Radial Velocity** and **Distance**.

All **type Ia supernovae peak at the same brightness** or absolute magnitude. This is because the progenitor stars of these events explode upon reaching the **limiting mass of 1.44 solar masses**. This is known as the Chandrasekhar limit and is consistent across all type Ia supernovae (see the 'An Introduction to Supernovae' worksheet for further information).

However, on Earth, we see different supernova targets vary in brightness, this is because they all lie at different distances away from us. We call this brightness their apparent magnitude (See the 'Calculating Magnitudes' worksheet for further explanation of these terms).

If we have values for both the apparent magnitude and absolute magnitude, we can calculate how far away the supernova is by using Equation 1 below.

Equation 1 – Distance

Where:

M = absolute magnitude m = apparent magnitude D = distance (pc)

*that's 10 to the power of the equation in brackets

This equation will give a value for distance in parsecs (pc), this is a unit often used in astronomy when using vast distances. One parsec is equal to 3.26 light years or 3.09×10^{13} km. This is equivalent to travelling to the Sun and back 103,000 times!

Even using this unit of distance, it is likely that you will still get very big values, therefore a more suitable unit to use is **megaparsecs (Mpc)**. This is equal to 1,000,000 (one million) parsecs and can therefore be calculated by dividing the number of parsecs by one million.

All of the values of apparent magnitude (m) for the supernovae targets are taken using a V-band filter on a telescope. **The absolute magnitude (M)** that all type Ia supernovae peak at in the **V-band** is equal to **-19.48**.

Open the '**Hubble Expansion Spreadsheet**' document and input the names of the supernovae and their corresponding peak magnitudes you have identified.

	H26	▼ (* <i>f</i> x							
A	A	В	С	D	E	F	G	H	
1		Supernova Name	Apparent Magnitude	Absolute Magnitude	Distance (pc)	Distance (Mpc)	Host Galaxy	Radial Velocity (km s^-1)	
3	Your Supernova			-19.48	78704.57897	0.078704579	F.		
4	1	SN2005fh	19.74	-19.48	698232404.1	698.2324041		42395	
5	2			-19.48	78704.57897	0.078704579	11:		
6	3			-19.48	78704.57897	0.078704579	11		
7	4			-19.48	78704.57897	0.078704579	6		
8	5			-19.48	78704.57897	0.078704579			
9	6			-19.48	78704.57897	0.078704579	11		
10	7			-19.48	78704.57897	0.078704579	11		
11	8			-19.48	78704.57897	0.078704579	6-		
12	9			-19.48	78704.57897	0.078704579			
13	10			-19.48	78704.57897	0.078704579			
14	11			-19.48	78704.57897	0.078704579	1		
15	12			-19.48	78704.57897	0.078704579			
16	13			-19.48	78704.57897	0.078704579			
17	14			-19.48	78704.57897	0.078704579			
18	15			-19.48	78704.57897	0.078704579	1		
19	16			-19.48	78704.57897	0.078704579			
20	17			-19.48	78704.57897	0.078704579			
21	18			-19.48	78704.57897	0.078704579			
22	19			-19.48	78704.57897	0.078704579	1		
23									

Figure 2 – The Hubble Expansion Spreadsheet

Obtaining the Radial Velocity of the Host Galaxies

Now we need to look up the **radial velocity** of the host galaxies of the supernovae. This value describes the speed at which a galaxy is moving towards or away from Earth. We want to know this value so that we can plot it against the distance of the supernovae and their galaxies to calculate how quickly the Universe is expanding.

Step 7: Refer to the '**Supernova Host Information'** worksheet which lists all the supernovae and the radial velocities of their associated host galaxies. Identify the radial velocity values for each of your targets and input these values into your table

Measuring the Expansion Rate of our Universe

Your table should now be complete. This means we can create a **Hubble plot** to find the rate that the Universe is expanding.

Step 8: To do this, plot your values onto the graph paper with '**Radial Velocity (km s**⁻¹)' on the y-axis and '**Distance (Mpc)**' on the x-axis.

Note: Make sure you use the distance column in values of Mpc and not pc.

Step 9: Next we want to add a **linear trendline** or a line of best fit. Estimate where this would go on your data.

Step 10: We refer to the rate at which the Universe is expanding as the **Hubble Constant**. This is equal to the **gradient** of your linear trendline and is in units of **km s⁻¹ Mpc⁻¹**. Use Equation 2 to calculate this gradient and the Universe expansion rate.

Equation 2 – Gradient of a Linear Line

Where:

y = point on y-axis m = gradient x = point on x-axis c = y-intercept

Hopefully you should have a value between **60 and 70 km s⁻¹ Mpc⁻¹**. This means that for every Mpc of space, the Universe is expanding at 60-70 km s⁻¹. That's around 30 times faster than a bullet!

Calculating the Age of the Universe

Now that we know how quickly the Universe is expanding we can calculate how old it is.

Step 11: Use the equation below to calculate the approximate age of the Universe:

Where:

t = approximate age of Universe (s) H_0 = your value of the Hubble Constant (km s⁻¹ Mpc⁻¹)

You will have got a very small number as your answer. This is because we need to convert the units of megaparsecs (Mpc) to kilometres (km) and the unit of seconds (s) to years (yrs).

Step 12: First convert your answer to km and record your answer below. This gives you the age of the Universe in seconds.

1 Mpc = 3.09 x 10¹⁹ km

Age in seconds = _____

Step 13: Now convert this answer to years and record your answer below.

1 yr = (60 x 60 x 24 x 365) s

Age in years = _____

You should have got an answer of around 14 billion years old!

Appendix 2 – Modified Julian Date

Julian Days refer to a continuous count of days since the beginning of the Julian period. Julian Day 0 therefore corresponds to noon on the 1st January, 4713 BC. Unlike the standard dating system, Julian Date is denoted by numbers and decimal fractions.

For any given moment since this beginning date, the Julian Date (JD) is denoted by the Julian Day number (beginning at Greenwich Mean Time noon), plus the fraction of the day at that particular moment.

The Modified Julian Date simplifies this notation. From the beginning of the Julian period up until now, Julian days have corresponded to between 2400000 and 2500000, MJD therefore eliminates this value from the number. MJD also begins at **midnight** rather than noon, therefore subtracting half a day from JD.

MJD is therefore derived from JD according to:

This modification means that MJD provides a 5-digit number for each day, followed by a decimal value corresponding to what time of day it is.

Astronomers commonly use this dating system as it provides a more straight forward method of performing chronological calculations.

Appendix 3 – Methods of Recruiting Schools and Teachers

Email

The first method used in an attempt to recruit teachers was sourcing their school email addresses. The My Local School website was used for schools in Wales and the government's Compare School Performance Service was used for schools in England. Both of these services allow you to search schools by region.

Given the researcher is based in Cardiff, Wales, a list of secondary schools in Central South Wales and South East Wales was searched. This returned 84 schools, (48 in CSW and 36 in SEW). These two regions were chosen as they were considered to be reasonable distances from Cardiff University. As participating schools required one or more visits, they needed to be within a reasonable commute for the researcher.

Each of these schools' website were then searched in order to source contact emails for the heads of science or science teachers where possible, otherwise for schools' administration office. Not all school websites provided contact information but those that did were all emailed with an invitation to the study and to use the educational resources. This invitation email can be found in Appendix 3.

As the researcher is originally from Oxfordshire and still has family living there, English schools within this area were also searched on the equivalent website. This returned 30 schools, and the same process was repeated.

All schools returned by these methods only included state schools and did not return any information for public schools.

Some email addresses were also already held of teachers who had previously attended CPD events ran by the Faulkes Telescope Project or had worked with them previously (mostly before the researcher's involvement). These teachers were also contacted via email and informed of these newly developed resources.

Physics Teacher Conferences

Every year the Institute of Physics (IOP) hosts a Welsh Physics Teachers' Conference in Wales. This invites teachers in Wales to a full day of continued professional development where they attend various workshops and speak to a number of exhibitors in the aim to take away materials and ideas they can use in their classroom.

The researcher attended the 2016 and 2017 events as an exhibitor where she was able to showcase the available resources and give teachers leaflets and postcards with more information and contact details. Teachers who showed interest were also able to 'sign-up' by providing their school and contact email.

After the event, these teachers were then contacted with an email similar to that mentioned previously, inviting them to participate in the study.

TeachMeets

TeachMeets are usually subject specific events held at schools. They provide teachers and practitioners an opportunity to discuss experiences and new teaching ideas in an informal setting.

The researcher attended several science-specific TeachMeets in South Wales where she was able to exhibit the resources and speak to teachers about applying them to their classroom. As in the IOP teachers' conference, teacher contact details were collected in order to follow up on any interest that was received.

Education Achievement Service Science Advisor

The researcher was fortunate enough to be put in contact with the Science Advisor for the Education Achievement Service (EAS) for South East Wales. The service is designed to promote education standards within Blaenau Gwent, Caerphilly, Monmouthshire, Newport and Torfaen (Education Achievement Service, no date).

The Science Advisor (SA) is responsible for going into schools within these five authorities in order to support their science departments. Through meeting the SA and demonstrating the resources on offer and the nature of this research, he was keen to get schools on board and involved in the study. This opened many doors as the SA was able to organise meetings with science teachers across many schools and led to many successful implementations.

Cardiff University Physics and Astronomy Outreach Blog

The Physics and Astronomy department at Cardiff University has its own online outreach blog. Here, posts are regularly published regarding things such as summarising articles of research within the department, news regarding various outreach events, and general physics and astronomy news.

This webpages draws attention from many local secondary schools, who are gearing students up to possibly go on to further their education at Cardiff University. This therefore presented another medium to reach out to local science teachers, promoting the study.

A blog was posted, that followed a similar format to the invitation email sent out to teachers with known email addresses and can be accessed online (Bartlett, 2017).

This prompted several responses from teachers who were interested in using the resources and taking part in the study.

Word-of-Mouth

As the study progressed, positive feedback was received from teachers. Many asked if they could use the material again and some even suggested other teachers that they thought would be interested. This word-of-mouth approach from teacher-to-teacher was effective. Presumably because the teachers feel more confident if they know something has been tested and trialled in a classroom and recommended by another teacher.

Other Access Routes

Several other means of reaching teachers were used on a smaller scale. The researcher and other members of the Faulkes Telescope project have worked closely with the PGCE (postgraduate certificate in education) science course at Cardiff Metropolitan University. They were able to provide contacts for previous students that are now in schools local to Cardiff. However, none of these teachers responded to emails.

Summary of Methods

Table 22 summarises the individual methods that successfully recruited each of the 11 participant groups in this study.

Participant Group	Recruitment Method
Pink	Email
Green	Word of Mouth
Blue	Education Achievement Service Science Advisor
Navy	CPD
Brown	Email
Yellow	TeachMeet
Orange	Cardiff University Physics and Astronomy Outreach Blog
Red	Cardiff University Physics and Astronomy Outreach Blog
Black	Email
Purple	Education Achievement Service Science Advisor
Grey	Education Achievement Service Science Advisor

Table 22 – Successful Recruitment Methods for Individual Participant Groups
Appendix 4 – Invitation Email to Teachers

Dear [Teacher Name],

I hope this finds you well and you don't mind me contacting you. My name is Sophie Bartlett and I am a postgraduate research student Cardiff University, looking into ways of engaging 14-18 year olds with their physics curriculum. I hope that with a context of astronomy and space, students can actively engage with some of their physics syllabus and enjoying their learning experiences.

I would like to offer you some free educational resources that have been newly developed by the Faulkes Telescope and Down2Earth projects at Cardiff University. All of our resources are designed to follow areas of the physics curriculum but apply astronomy and space science as a context. The materials are suitable for 14-18 year olds and the activities are designed to follow processes of inquiry that allow students to develop their scientific skills.

We have two main "Teacher Packages". The first is our Hubble's Expansion of the Universe project, found here: <u>http://resources.faulkes-telescope.com/course/view.php?id=143</u>. This material involves using observations of supernovae (taken with the Faulkes Telescopes) to calculate how quickly the Universe is expanding and how old it is using Hubble's Law. Students are able to analyse pre-packaged data sets or can even open their own telescope account with us to perform their own observations.

The second is our Down2Earth material that can be freely downloaded here: <u>https://drive.google.com/drive/folders/0B73hLNgTXoTuOENpQzI4dENhc2M</u> These activities use an online simulator that allows students to investigate various comet and asteroid impact events and their effects on Earth. Students are able to explore topics such as speed, distance and time, kinematics, and energy.

As I am a research student, I am looking to investigate whether or not these materials can successfully engage students with areas of their physics syllabus. I am looking to do this through short questionnaires and focus groups and also by observing students whilst they use the material.

If you are interested in using the material, you will also be invited to take part in this study. Although participation is not compulsory in using the materials, it would be a great help to my studies. If you did choose to participate, please be aware that you are free to withdraw at any time without any consequence and may continue to use the materials as you wish.

If this is something you would be interested in, please get in touch and I will be happy to explain things in more detail.

I look forward to hearing from you. Warm regards, Sophie

Appendix 5 – Ethical Approval Form

Please tick one box only: □ Staff ■ PhD / MPhil

Title of Project: The Impact of Astronomy-based Educational Resources on Students' Engagement in their Physics Syllabus

Name of Student: Sophie Bartlett

Name of Supervisor: Paul Roche / James Snook

Contact email address: bartletts2@cardiff.ac.uk

Date: 23/03/2017

Participants		yes	no	n/a
Does the research	Children (under 16 years of age)	Х		
involve participants	People with learning difficulties		Х	
following groups?	Patients or relatives or carers of patients (NHS approval is required)		Х	
	People in custody		Х	
	People engaged in illegal activities		Х	
	Vulnerable elderly people		Х	
	Any other vulnerable group		Х	
When working with children: I have read the interim guidance for researchers working with children and young people		Х		

Consent Procedure	yes	no	n/a
Will you describe the research process to participants in advance, so that they are informed about what to expect?	X		
Will you tell participants that their participation is voluntary?	X		
Will you tell participants that they may withdraw from the research at any time and for any reason?	X		
Will you obtain valid consent from participants (specify how consent will be obtained in box A) ²	X		
Will you give participants the option of omitting questions they do not want to answer?	X		
If the research is observational, will you ask participants for their consent to being observed?	X		
If the research involves photography or other audio-visual recording, will you ask the participants for their consent to being photographed / recorded and for its use / publication?			Х

Possible Harm to Participants	yes	no	n/a
Is there any realistic risk of any participants experiencing either physical or psychological distress or discomfort?		Х	

² if any non-anonymous and/or personalised data will be generated or stored, written consent is required.

Is there any realistic risk of any participants experiencing a detriment to their interests as a	Х	
result of participation?		

Data Protection			no	n/a
Will any non-anonymous and/	/ill any non-anonymous and/or personalised data be generated or stored			
If the research involves non-	Gain written consent from participants?			Х
personalised data, will you:	Allow participants the option of anonymity for all or part of the information they provide?			Х

Research Governance	yes	no	n/a
Does your study involve the use of a drug?		Х	
You need to contact Research Governance before submission (resgov@cf.ac.uk)			
Does your study involve the collection or use of human tissue?		Х	
You need to contact the Human Tissue Act team before submission (<u>hta@cf.ac.uk</u>) Advice can be			
sought from the College representative: Prof Adrian Porch (PorchA@cardiff.ac.uk)			

If any of the shaded boxes have been ticked you must explain in Box A how the ethical issues are addressed. The list of ethical issues on this form is not exhaustive; if you are aware of any other ethical issues you need to make the SREC aware of them.

Box A The Project (provide all the information listed in a separate attachment)

The study focuses on evaluating new educational resources that apply an astronomy context in order to teach parts of the UK's national curriculum for physics in key stages 3, 4 and 5.

These resources will apply an inquiry-based pedagogy and comprise of investigative activities using real scientific data obtained from telescopes and satellites and also computer simulations.

The resources will be evaluated in terms of their impact on students' engagement and enjoyment in physics. Pre- and post-questionnaires have been developed for students to complete in order for comparisons to be made before and after students have used the resources. Feedback questionnaires will also be given to teachers in order to assess if their opinions reflect those expressed by their students.

Classroom observations and (where possible) focus groups will also be carried out in order to expand on the quantitative findings and to develop a deeper understanding of students' engagement with physics in their classroom, the educational resources and also their learning experiences as a whole.

No sensitive data will be collected from participants and all data is anonymous and nonidentifiable.

Prior to any data collection, all participants (students and teachers) will be given an overview of what the study entails, its aims and objectives and methods of data collection. In instances where the students are under the age of 16, this will be provided to the parents. All participants will be notified that all data that is collected is anonymous

and securely stored.

They will also be provided with an opt-out form should they wish to not take part. In instances where the participants are under the age of 16, parents will be able to opt-out on behalf of their child should they wish to do so. Consent must also be provided by the teacher in order for the study to pursue.

All participants will be made aware of their right to withdraw from the study at any time without consequence, and any data collected on them will be destroyed. They will be clearly informed on how they can withdraw from the study if they wish to do so.

Any participants who opt-out of the research will not have any information stored and will not be included in the research and data analysis, however they will still be present within the research environment (their classroom). Questionnaire and focus group data will not be collected from these participants and although present, they will be excluded from the data collection in the classroom observations.

Researcher's declaration (tick as appropriate)

I consider this research project to have negligible ethical implications (can only be used if none of the shaded areas of the checklist are ticked)

I consider this project to have some ethical implications. Box A clearly describes the ethical issues and how they are addressed.

I consider the project to have significant ethical implications. Box A clearly describes the ethical issues and how they are addressed.

Researcher Signature:

Name: Sophie Bartlett

Name:

Date: 23/03/2017

StBarllett

Principal Investigator or Supervisor

Signature:

Dr. Paul Roche

Date: 23/03/2017

Appendix 6 – Ethical Approval

School of Physics and Astronomy Cardiff University Queen's Building The Parade Cardiff CF24 3AA

24 April 2017

Dear Ms Bartlett;

I am writing to you in my capacity as the Ethics Officer for the School of Physics and Astronomy with respect to your application for ethics approval of your student project entitled "The Impact of Astronomy-based Educational Resources on Students' Engagement in their Physics Syllabus".

I am pleased to inform you that following due consideration by the School Research Ethics Committee your application for ethics approval has been accepted.

Please note that projects that require ethical approval will be monitored on an annual basis by the School Ethics Officer to ensure that agreed standards are being met. Researchers will be required to report on projects and provide evidence of the research methods adopted as appropriate.

Please do not hesitate to contact me if you require any further information. I can be contacted either by email at www.lewisR54@cardiff.ac.uk or alternatively on extension 75433.

Yours sincerely,

nh-s

Dr Richard Lewis CPhys MInstP FHEA School Ethics Officer

Appendix 7 – Information Sheets and Consent Forms

Information Sheet for Teachers

Research Project into Astronomy and Physics Educational Materials

Invitation

You are invited to participate in a research study ran by the Faulkes Telescope Projects. The project provides a set of educational materials based around themes of astronomy and space science whilst following some of the topics in the physics curriculum.

We want to investigate the impact these materials have on students' engagement and enjoyment in physics. To do this we would like to ask you and your students some questions before and after you implement the materials in your classroom and watch how your students interact with them. We would also like to invite some students to participate in a short focus group.

The study is being conducted by Sophie Bartlett, a student at Cardiff University who is researching physics education in Wales and England. Before you decide on whether or not you would like both yourself and your class to participate in the study, it is important that you understand what the research involves and why it is being carried out.

Please take the time to read through the information provided below. If you have any further queries please do not hesitate to contact Sophie via the contact details provided in this document.

If you wish to participate in the study, please complete and sign the attached consent form. If we do not receive the form, we will assume you and your students do not wish to participate and the study will not pursue.

What are the astronomy and physics educational materials?

There are a variety of educational materials involving themes of asteroids, comets and impacts, the life cycle of stars and the expansion of the Universe. Some of them involve an online computer simulation that allows you to create your own impactors and investigate the consequences that would result from their impact with Earth. Some involve collecting and analysing real astronomical data and using it to measure our Universe. They are designed to follow the contents of the physics curriculum that you follow in your usual physics or science lessons.

What is the research?

In this study, we are looking to explore how your students interact with the educational materials that we provide to your school. We want to find out whether or not they find the materials enjoyable and useful to their learning of the physics syllabus and what improvements can be made to them. We also want to explore how you think your students respond to the material, how useful you perceive them to be to their learning and what improvements you think could be made.

Who is being invited to participate and what will it involve?

We are looking for physics teachers in Wales and England and their students to participate in the study. If you consent to participation then we will then invite your individual students to participate. If your students are under the age of 16 then we will also provide an **opt-out** consent procedure to their parents. If you do chose to participate, it will involve you implementing one or some of the educational materials in your class and completing before and after questionnaires relating to your students' education in physics and the materials themselves. The researcher would also like to be present when you implement the materials in your classroom. This is in order to make some observations about how your students interact with the materials. It is possible you may also be invited to take part in a short interview where you will be asked a few more questions about your experiences of using the materials and upon your consent, will be voice recorded in order to help with the analysis.

Are there any risks or benefits to me or my students from taking part in this study?

There are no foreseeable risks associated with this study other than the time element, however all the materials are designed to coincide with the physics syllabus that would be covered in usual physics lessons.

What happens to the information me and my students provide?

All the information that you provide will only be seen by the researcher and will not be seen by anyone else, this includes students and their parents. All information will be de-identified and will not be recognisable to you personally. This also goes for all the information that your students provide.

The information that you and your students provide will be used in a PhD thesis and future publications about physics education. The thesis and publications will be accessible to the public but all information will be anonymised and not identifiable.

What if I or my students don't want to take part?

Participation in this study is optional and you do not have to partake. If you do not wish to participate, simply do not return the attached consent form. You do not have to give a reason for your decisions and it will not disadvantage you or your students in anyway. Students are also free to decline their participation even if you consent and they are also free to withdraw at any point during the study even if you wish to continue.

Even if you do not wish to take part in the study you will still be able to use the educational materials with your class without evaluation should you wish to do so.

What if I change my mind about participation?

If you consent to participation and change your mind during the study you are free to withdraw at any point. Any information gathered before this point will be removed from the research data. You do not need to provide a reason for withdrawal and you and your students will not be disadvantaged as a result in any way.

Confidentiality

All information that is gathered by the researcher will be securely stored and will only be accessible to the researcher, except as required by law. No data that is identifiable to you will be made public. This is the same for the information provided by your students.

Data files will be stored for up to 5 years to coincide with the protocols of Cardiff University. These will be kept secure on a password protected computer system or in a locked filing cabinet.

Contact Information

If you would like more information about the research or have any queries, please do not hesitate to contact the researcher:

Sophie Bartlett Phone: 07794 564 249 Email: <u>BartlettS2@cardiff.ac.uk</u>

Thank you for considering this invitation. This information sheet is for you to keep. Please refer to the next page for the consent form

Consent Form for Teachers

Research Project into Astronomy and Physics Educational Materials

Name of Researcher: Sophie Bartlett

I confirm that I have read and understood the enclosed information sheets and have had any queries answered to my satisfaction.

I give consent freely for my participation in the above study and I understand that both my students and myself are free to withdraw from the study at any time, without it disadvantaging me or them, or preventing us from using the educational materials. I understand that both my students and myself are free to withdraw from the study at any time and it is not necessary to provide any reason for this.

I understand that mine and my students' personal information will be de-identified and remain anonymous.

I understand that this research will be published in a PhD thesis and may be later published in peer-reviewed educational journals but all information will be anonymous and unidentifiable.

I consent to my students and myself (please tick if you agree):



Completing a questionnaire on two occasions

Taking part in an interview

Taking part in a classroom observation

Print Name:	

Signature: _____

Date: _____

Appendix 8 – Students' Perceptions of Physics Questionnaire (Pre) STUDENT PRE-QUESTIONNAIRE

We are interested to hear about your thoughts towards <u>physics in school</u>. This is not a test. There are no correct or incorrect answers, we merely want to hear your own personal opinions.

These questions ask about your <u>physics lessons</u>. You should base your answers on your physics lessons for <u>this year</u> and not ones you have had in the past (e.g. in previous school years).

Please also make sure you have provided your date of birth and gender below.

Date of Birth:

Gender:

1/. What is your FAVOURITE subject at school?

2/. What subject do you think you are <u>BEST</u> at?

3/. Using the scale provided, please indicate the extent to which you agree with the following statements.

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
We learn interesting things in physics lessons	1	2	3	4	5
I look forward to my physics lessons	1	2	3	4	5
Physics lessons are exciting	1	2	3	4	5
I enjoy my physics lessons	1	2	3	4	5
Physics lessons increase my interest in physics topics	1	2	3	4	5
I wish we had more physics lessons	1	2	3	4	5
Physics lessons are boring	1	2	3	4	5

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
I am good at physics	1	2	3	4	5
I get good marks in physics	1	2	3	4	5
I learn physics quickly	1	2	3	4	5
I understand what we learn in physics lessons	1	2	3	4	5
Physics lessons are useful to my learning	1	2	3	4	5
I find physics difficult	1	2	3	4	5
I struggle in my physics lessons	1	2	3	4	5

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
I would like to study more physics in the future	1	2	3	4	5
I would like to study physics in A- level	1	2	3	4	5
I would like to study physics at university	1	2	3	4	5
I would like to have a job working with physics	1	2	3	4	5
I would like to become a scientist or physicist	1	2	3	4	5
Physics is useful for my future	1	2	3	4	5
I would like to do research in physics	1	2	3	4	5

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
I have a strong interest in learning about the Universe	1	2	3	4	5

I find astronomy and space science interesting	1	2	3	4	5
I am interested in what is currently happening in the field of astronomy and space science	1	2	3	4	5
Being involved in astronomy and space science research is exciting	1	2	3	4	5
Astronomy and space science is an exciting hobby	1	2	3	4	5
I would like to study more astronomy and space science in the future	1	2	3	4	5
I would like to study astronomy and space science at university	1	2	3	4	5
I would like to have a job working with astronomy and space science	1	2	3	4	5

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
Investigation work in my physics lessons is interesting	1	2	3	4	5
I learn physics better when we do investigation work	1	2	3	4	5
I would rather do a physics investigation than <u>read</u> about physics	1	2	3	4	5
I would rather do a physics investigation than be <u>told</u> about physics	1	2	3	4	5
We often carry out investigations in physics lessons	1	2	3	4	5
We often design investigations ourselves in physics lessons	1	2	3	4	5
I dislike investigation work in physics	1	2	3	4	5

In physics lessons I am able to...

	Never	Rarely	Sometimes	Often	Always
Discuss events in the news that					
are related to what I am	1	2	3	4	5
learning in physics lessons					

to my life outside of school	1	2	3	4	5
See how what I learn in physics					
is relevant to my day-to-day	1	2	3	4	5
life					
Apply my everyday					
experiences to my physics	1	2	3	4	5
lessons					
Apply what I already know to	1	2	2	Δ	E
my work in physics lessons	L L	Z	5	4	5
Relate what I learn in physics	1	2	2	4	-
to the world around me		2	3	4	5

4/. What do you like MOST about physics lessons?

Relate what I learn in physics

5/. What do you like <u>LEAST</u> about physics lessons?

6/. Where 1 = <u>not at all confident</u> and 10 = <u>extremely confident</u>, how confident do you feel in your <u>understanding</u> of the following topics?

Astronomy and space science	1	2	3	4	5	6	7	8	9	10
General physics	1	2	3	4	5	6	7	8	9	10
General mathematics	1	2	3	4	5	6	7	8	9	10
The scientific method	1	2	3	4	5	6	7	8	9	10

7/. Do you have any other comments about your physics lessons?

Thank you for taking the time to complete this questionnaire.

Appendix 9 – Students Perceptions of Physics Questionnaire (Post)

STUDENT POST-QUESTIONNAIRE

We are interested to hear about your thoughts towards the <u>activity(s) you have completed</u>. This is not a test. There are no correct or incorrect answers, we merely want to hear your personal opinions.

Please also make sure you have provided your <u>date of birth</u> and <u>gender</u> below.

I learnt how to do the activity quickly	1	2	3	4	5
I understood what we learnt in the activity	1	2	3	4	5
The activity was useful to my learning	1	2	3	4	5
I found the activity difficult	1	2	3	4	5
I struggled with the activity	1	2	3	4	5

Date of Birth:

Gender:

1/. Using the scale provided, please indicate the extent to which you agree with the following statements.

When we carried out the activity...

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
We learnt interesting things in the physics lesson	1	2	3	4	5
I would look forward to doing more physics lessons like this	1	2	3	4	5
The physics lesson was exciting	1	2	3	4	5
I enjoyed the physics lesson	1	2	3	4	5
The physics lesson increased my interest in physics topics	1	2	3	4	5
I wish we did more physics lessons like this	1	2	3	4	5
The physics lesson was boring	1	2	3	4	5

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
I was good at doing the activity	1	2	3	4	5
I think I will get good marks in the activity	1	2	3	4	5

After doing the activity...

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
I have a strong interest in learning about the Universe	1	2	3	4	5
I find astronomy and space science interesting	1	2	3	4	5
I am interested in what is currently happening in the field of astronomy and space science	1	2	3	4	5
I think being involved in astronomy and space science research is exciting	1	2	3	4	5
I think astronomy and space science is an exciting hobby	1	2	3	4	5
I would like to study more astronomy and space science in the future	1	2	3	4	5
I would like to study astronomy and space science at university	1	2	3	4	5
I would like to have a job working with astronomy and space science	1	2	3	4	5

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
The investigation work in the activity was interesting	1	2	3	4	5

I learnt physics better when we did the activity	1	2	3	4	5
I would rather do an activity like this than <u>read</u> about physics	1	2	3	4	5
I would rather do an activity like this than be <u>told</u> about physics	1	2	3	4	5
We got to carry out an investigation in the activity	1	2	3	4	5
We got to design an investigation for ourselves in the activity	1	2	3	4	5
I disliked the investigation in the activity	1	2	3	4	5

In the activity, I was able to...

	Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
Discuss events in the news that are related to the topic of the activity	1	2	3	4	5
Relate what I learnt in the activity to my life outside of school	1	2	3	4	5
See how what I learnt in the activity is relevant to my day-to-day life	1	2	3	4	5
Apply my everyday experiences to the activity	1	2	3	4	5
Apply what I already know to the work in the activity	1	2	3	4	5
Relate what I learnt in the activity to the world around me	1	2	3	4	5

4/. Where 1 = <u>not at all confident</u> and 10 = <u>extremely confident</u>, how confident do you feel in your <u>understanding</u> of the following topics after completing the activity?

Astronomy and space science	1	2	3	4	5	6	7	8	9	10
General physics	1	2	3	4	5	6	7	8	9	10
General mathematics	1	2	3	4	5	6	7	8	9	10
The scientific method	1	2	3	4	5	6	7	8	9	10

5/. Did you learn anything new in the following topics?

Astronomy and space science	Yes / No
General physics	Yes / No
General mathematics	Yes / No
The scientific method	Yes / No

6/. If you answered yes to any of Question 5, please provide some details on what you learnt.

2/. What did you like MOST about the activity?

7/. How do you think the activity could be improved?

8/. Do you have any other comments about the activity?

3/. What did you like LEAST about the activity?

Thank you for taking the time to complete this questionnaire.

Appendix 10 – Classroom Observation Demographics and Protocol

Date:						
School:			Period/Time:			
Activity:						
Participants:	Male:			Female:		
Gender of Teacher: Support Staff:						
Class:			Year Gro	up:		
Seating Plan (drawn) Implemented By:						
Working Arrangements: Group	Individual		/	Pair	/	
Computer Use:						
Noise Levels:	High	/		Moderate	/	Low

Delivery and Structure of Lesson	
Event	Duration
Pre-Planning and Preparation	
•	
Students provided with:	
•	
Delivery of Activity	
•	
Structure of Lesson	
•	

Unplanned Activities	Individual, Paired or Group Work
This was to detail any actions that students performed that was not asked of them by the teacher or was not in the scope of the activity. This could have been a positive or negative behaviour.	This covers how students were working. If they are in pairs or groups, are there specific roles? Is there a clear leader? How are roles determined? Etc.
Capability of Set Tasks	Communications
How students coped with the initiative activity. Did they encounter any problems? If so, how were they overcome? Did they seek help from the teacher or peers or did this cause disengagement? It also looked at students' approaches to tasks as well to explore how they investigated a problem or question.	What is communication like among students and their peers and the students and the teacher? Do students work individually in silence? Do they call out or must they raise their hands? Is communication between students and the teacher formal or informal? Are all communications task focused or does conversation stray to other, irrelevant or relevant topics.
Time Optimisation	Body Language
How is lesson time spent? What is time attributed to? What everything achieved that was set out at the beginning of the lesson? Are all students working at the same pace? Are there extension tasks for those students finishing ahead of others?	Body language and kinesics can reveal a lot about the attitude of an individual to a task or topic. This area therefore focused on non- verbal behaviour of students.
Behaviour of Teacher	Behaviour of Students
What was the teacher's role in the lesson? Were they seen to facilitate learning or do they employ a didactic approach, prescribing students with specific instructions? Do they walk around the classroom and interact with students or do they remain at the front?	This is perhaps the broadest area of the protocol and is intended to encompass students' behaviour in terms of specific physical actions. This would focus on aspects such as noise levels, focus to the task and how long it takes students to focus and work on the task at hand. How do they react to the OA? Are students independent or are there levels of conformity in the classroom?
	1

Comments from Students

These are comments made by the students throughout their lesson. These comments could be to each other, the teacher or to the researcher. Comments of interest are ones related to what they are doing in the lesson, the teacher or anything relevant to that particular lesson. Comments made that are entirely irrelevant to the lesson are also of interest as they could be indications of disengagement with the task.

(also say if comments were prompted in any way)

Comments from Teacher

This section focuses on particular quotes from teachers with relevance to the students, the lesson or the activity. These comments could be directed at the researcher or at the students and could be positive or negative.

(also say if comments were prompted in any way)

Researcher's Interpretation (for after completion of observation only)

Appendix 11 – Student Focus Groups

Group Navy:

The focus group carried out in Group Navy involved nine participants out of the 26 that participated in the activity. The focus group was carried out in the students' usual physics classroom during a lunchtime, 17 months after implementation of the OA. The question schedule was as follows:

- 1. How would you describe your day-to-day physics lessons?
- 2. Are you able to name three good things and three bad things about your normal physics lessons?
- 3. Do you remember using the Impact Calculator? *Prompt: show them the simulator*
- 4. How would you describe the activity that used the Impact Calculator?
- 5. Can you name three good things and three bad things about the activity?
- 6. Do you feel like you learnt anything new in the activity?
- 7. Do you feel you applied any skills?
- 8. What did you think of the topic of space and impacts?
- 9. Was there anything you didn't like about the Impact Calculator?
- 10. Would you be interested in doing more activities like this?
- 11. Is anyone interested in pursuing physics at university?

Questions for Group Brown:

The focus group with Group Brown took place two months after implementation of the OA with six of the 25 students who completed the activity. The focus group was performed in a break out room during lesson time. The question schedule was as follows:

- 1. How would you describe your normal physics lessons? *Prompts: Do you enjoy them? What sort of activities do you do in lessons? How do the lessons run?*
- 2. Why would you say is in control of your learning in physics lessons? Why do you think that?
- 3. How confident do you feel in your ability in physics?
- 4. Can you name three good things and three bad things about your normal physics lessons?
- 5. Do you remember when I was here before, you used the Impact Calculator to do an activity?

Prompts: show them Impact calculator and activity handouts

- 6. Do you feel like you understood what you had to do in the activity?
- 7. Can you name three good things and three bad things about the activity?
- 8. Were you able to take anything away from the activity?
- 9. Who was in control of your learning during the activity?
- 10. How did your confidence when doing the activity compare to your confidence in physics lessons? Was it the same or different?
- 11. Would you like to do activities like this in all your lessons?
- 12. How do you feel you work best in science? Individually or in groups?

Appendix 12 – Teacher Interview Questions

Interview with Mr Brown

The interview with Mr Brown was carried out one month after implementation of the OA and took place in his classroom at the end of a school day. The school had invited the researcher to give a talk to a group of year 11s so the researcher stayed at the school afterwards in order to conduct the interview. The questions asked were as follows:

- 1. How do you think the students performed in the activity?
- 2. Do you think the majority of students were able to complete the task? What gave you that impression?
- 3. Do you think that even though the simulator was engaging it wasn't necessarily appropriate as it may have detracted too much from the lesson content?
- 4. Student [M1], you spoke very highly of this student during my first observation lesson, what do you think of that student?
- 5. Are there any lessons where he doesn't necessarily work as hard or perform? Are there any circumstances in particular that you find he doesn't connect with?
- 6. Student [Isabel], you also spoke highly of this student, what is your thoughts on this student?
- 7. And is she hard working?
- 8. Can you describe the types of students that you think would benefit from this activity? Like is there a particular group would find the lesson particularly beneficial?
- 9. Okay so my next question was what students do you think wouldn't benefit from the activity?
- 10. Okay so if you were to run the activity again with those students, how would you run it differently?
- 11. Do you think the level of the content was suitable for the students you ran it with?
- 12. Students had to report their confidence in their questionnaires. Many students' confidence was higher for the activity than in usual physics? What do you think could be the reason for this?
- 13. Do you think it was too easy for them?
- 14. Okay so it was more of like a revision activity?
- 15. Would you use this sort of activity in every lesson?
- 16. Why?
- 17. Okay, maybe not using simulations in every lesson but would you do every lesson as a process of investigation and using inquiry?
- 18. Okay so what would be the realistic amount of lessons you could incorporate it into?
- 19. So if one existed, so ideally you would implement them in every lesson, so how many would you implement if they were readily available?
- 20. Do you think that the resource was accessible for all the students in your class?
- 21. Before implementation all teachers are told how the activity can be differentiated would you rather be provided with the most involved version (hardest) so that you could add in scaffolds, or the simplest version?
- 22. Yes, so would you rather I gave you a resource that was the bare structure and you could add in various scaffolds...
- 23. So the hardest version of this activity would be for example, you need to investigate how changing the projectile diameter influences the kinetic energy and the size of the crater. And then they'd have to come up with a hypothesis, how they're going to investigate it and a table to record it. But for your students, we gave them the process and a table to record.
- 24. What made you decide to not have them working in groups?

- 25. So because you gave everyone a number according to what parameter they were investigating, that meant everyone on each table was investigating something different so they couldn't work together to investigate the same thing
- 26. So do your class benefit more from individual work than group work?

Interview with Mrs Yellow

The interview with Mrs Yellow took place immediately after her second implementation of the OA, thus approximately one year after her first implementation. As the implementation took place in a final lesson period which the researcher observed, the interview was carried out immediately after, at the end of the school day, in Mrs Yellow's classroom. The questions asked were as follows:

- 1. First of all what did you think of the activity overall?
- 2. How do you think the students performed in the activity?
- 3. Do you think the majority of students were able to complete the task?
- 4. And what gave you that impression?
- 5. Do you think the activity covered a suitable amount of curriculum content?
- 6. Do you think the objective/goals were clear from the activity?
- 7. Can you describe the types of students that you think would benefit from this activity?
- 8. Are there any students that you feel wouldn't benefit?
- 9. Students had to report their confidence in their questionnaires. Many students' confidence was higher for the activity than in usual physics? What do you think could be the reason for this?
- 10. Would you want to use this sort of activity in every lesson? That involves students looking at real data?
- 11. Do you think that the resource was accessible for all the students in your class?
- 12. There are several ways of how this activity can be implemented and also differentiated would you rather be provided with the highest level (hardest) so that you can add in guidance or would you rather have an activity that provided all guidance (simple) and you could take out bits?

Appendix 13 – Examples of Thematic Analysis of Observational Data

Table 23 – Examples of Thematic Analysis of Observational Data			
Student Confidence			
POSITIVE	NECATIVE		
Girl 1 in group B said to girl next to her "so I can do	Students other than 1 and 2 often wouldn't ask		
science"	many questions – even if they didn't understand something – they would typically wait for the		
F2 very intelligent and knowingly so – often answered without raising hand – other females	teacher to notice they were struggling or just be told what to do.		
around her actually left her to answer questions, even if I was looking at them and directing a	When teacher told them I had said they'd performed		
question at them, they would look to her to answer.	better than private school students – they cheered. At the end 2 jokily said "were we really better or		
2 – Although not the most able in the group (but expected as was a yr 10 among year 12s and year 13s) – was very confident in giving it a go and asking questions to researcher – not afraid to give everything ago. Promising example of where a	were you just saying that to make us feel better?" – Almost still seeking approval or confirmation. Perhaps suggesting a lack of confidence?		
student could be potentially intimidated by their peers and activity but embraced the task and shone.			
Female participant had been unsure throughout activity. Researcher showed her answer sheet at end to prove she had done it correctly. Responded with "I should have done physics"			
Student E	ngagement		
More focused behaviour than in observation of	Once called out their readings in collaborative task		
standard, day-to-day physics lesson. Even	- became slightly unsettled and talking among		
conversations among each other were more task- based this time whereas first lesson was more just	themselves – suggested bare minimum – will do the work but not particularly interested		
chatting (particularly girls)			
	Distractions – talking about unrelated topics at		
Students didn't pack up immediately at the end of the lesson – wanted to finish what they were doing	times		
- interpretation was unclear, could indicate	(standard observation) Girls didn't show such quick		
engagement but could also have been because they	turnaround as boys did- once they were chatting		
didn't want to finish work for homework.	among friends it was harder for teacher to rein them		
Multiple occasions where students asked questions	in and rocus on the task at hand.		
related to but beyond scope of activity – indicative			
of genuine interest?			
Inquiry-Based Learning			
Group $E = 5$ had a teaching assistant- very, very engaged, though this could have been influenced by	Some students were playing around with random parameters – although not directly doing the set		
to be thinking out loud a lot – beyond scope of	investigation what happened when they changed		
activity and coming up with interesting theories and	things.		
questions for further exploration.			
Group A – this group realised they didn't really	Some of the girls couldn't find the data they needed to ask the researcher or teacher. Boys however		
know how the different materials of the impactor	seemed to explore the impact calculator interface		
differed from one another in terms of density. So	until they found the source of this data		
explore what happened when this is the only			
variable that they changed and even produced a			
graph to plot their results. This was interpreted that			

Note: Extracts are colour-coordinated according to group.

 their curiosity was triggered and they had an interest in the subject matter. They took it upon themselves to do extra work that wasn't necessary in order to explore things further. Not only was this additional work – it was something they had come up with themselves, not suggested by teacher. Group D - All playing around with the impact calculator – seeing how they could wipe out a whole country – talking to each other about the activity and what results they were getting right from the offset 	
Student-Teacl	er Interactions
Girls showed noticeably more communication with the researcher and teacher both in comparison to the day-to-day lesson and compared to boys – asking questions. Boys less so, although there was a teaching assistant that tended to stick with the boys they weren't very communicative with him. Appeared teacher would consciously avoid telling students they were wrong – responses such as	 Group Orange teacher seemed to focus most of her attention on Hollie and James – the most able and most focused students. This was almost opposite in group Blue – teacher paid his most attention to the girls that were typically disengaged and didn't do the work. Joking with the girls on group B – saying he was shocked that they'd done all the work – to one girl
"you're nearly there", "good suggestion but not quite the answer I'm looking for"Relaxed classroom climate. Teacher had informal conversations with students, made jokes.	"are you feeling okay?" – obviously this was not typical behaviour for them to get all the work done in class "there is a state of shock and disbelief, they've actually done the work" "do you know what, I am hugely impressed with you guys, I've got to say"
Situation	al Interest
Male in group $6 -$ "sir, this is the most entertaining	One group noticeably not an apping with activity
 thing we've ever done" Comments from girl 1 in Group B (who were first in class to complete the activity) said – "I never do any work in this class, look at my book, there's pretty much nothing in it". Girl on the end in Group C also reacted to this and said "omg I've actually completed the work, I never do anything, look at my last less, I wrote one sentence", then "I actually did the work today" – showed research books and this was true – all lesson work is supposed to be recorded in their books. Ellie – "wow that's in seconds? [referring to age of 	one group noticeably not engaging with activity and having conversations that were not relevant to activity. Female student near beginning of activity "what are we doing?" and "what's the point in this?"
 thing we've ever done" Comments from girl 1 in Group B (who were first in class to complete the activity) said – "I never do any work in this class, look at my book, there's pretty much nothing in it". Girl on the end in Group C also reacted to this and said "omg I've actually completed the work, I never do anything, look at my last less, I wrote one sentence", then "I actually did the work today" – showed research books and this was true – all lesson work is supposed to be recorded in their books. Ellie – "wow that's in seconds? [referring to age of the universe] Well I've learnt something cool" 	The group noticeably not engaging with activity and having conversations that were not relevant to activity. Female student near beginning of activity "what are we doing?" and "what's the point in this?"
thing we've ever done" Comments from girl 1 in Group B (who were first in class to complete the activity) said – "I never do any work in this class, look at my book, there's pretty much nothing in it". Girl on the end in Group C also reacted to this and said "omg I've actually completed the work, I never do anything, look at my last less, I wrote one sentence", then "I actually did the work today" – showed research books and this was true – all lesson work is supposed to be recorded in their books. Ellie – "wow that's in seconds? [referring to age of the universe] Well I've learnt something cool" Student R	cone group not engaging with activity and having conversations that were not relevant to activity. Female student near beginning of activity "what are we doing?" and "what's the point in this?"

1

Appendix 14 – Examples of Content Analysis of Open PPQ Responses

1 able 24 - what do you like <i>most</i> about physics lessons?		
Theme	Response	
	"experiments and learning by doing it instead of being told about it"	
Practical Lessons	<i>"I enjoy doing practicals as I get more involved in the lessons and I find I learn more"</i>	
	"doing experiments"	
	"learning things that I did not know about before"	
Learning	<i>"learning new and interesting topics, such as space, gravity and momentum which give me insight into the world"</i>	
	"giving my opinion and learning"	
	"Universe, interesting to talk about"	
Astronomy	"learning about astronomy and why it's important and relevant"	
	"I enjoy learning about the Universe and potential changes to the Universe"	
	"the teacher"	
The Teacher	"my teacher"	
	"our teacher's fun and good at explaining things"	
	"learning new formulas"	
Mathematics	"working out calculations"	
	"working out equations"	
	"particle physics and quantum physics"	
Particle Physics	"particle physics"	
	"learning about new, abstract concepts e.g. quantum physics"	
	"variety in the subject"	
Diversity	"physics is a free subject! There are lots of different aspects to it"	
	"variety of topics – from mechanics to the standard model"	
	"you almost always learn something new that's really interesting"	
Interesting	"topics that I find interesting and important"	
	<i>"learning new interesting topics that give me an insight into how the world works"</i>	

 Table 24 - What do you like most about physics lessons?

1	able 25 - What do you like <i>least</i> about physics lessons?
Theme	Response
	"writing lots"
Writing	"written questions"
	"all the writing"
	"boring presentations"
Boring	"boring facts and having to write them down in our books"
	"when it isn't interesting, it's extremely boring"
	"the need for technical language"
Difficulty	"I don't find it easy to learn and I struggle very much because we hop from one subject to another"
	"the fact that we constantly used advanced vocabulary, it makes me feel nervous"
	"electricity"
Specific Topics	"circuits"
	"space"
	"required praticals"
Practical Lessons	"that we don't do many practicals or investigations"
	<i>"testing practicals which I view as a little pointless because you could easily find the answer on line or in a book"</i>
	"reading and using textbooks"
Reading	"lots of reading"
	"reading tasks"
	"plotting graphs"
Working with Graphs	"graphs"
	"graphs and tables"
	"equations"
Mathematics	"having to learn lots of equations"
	"the equations and maths"
	"teachers show no interest or enthusiasm in presenting the work"
The Teacher	"long talks by the teachers"
	"just sitting in class and listen to what the teachers say for around an hour"

Table 25 - What do you like *least* about physics lessons?

Tuble 20 What and you me most about the activity. (Down2Darth)			
Theme	Responses		
	"using the simulator"		
The Simulator / Impact Calculator	"the impact calculator"		
	"using the app on the iPad to relate to the topic"		
	"the investigations part"		
Investigation	"how 'hands on' it was"		
	"seeing how it changed by changing something"		
	"I enjoyed looking at how asteroids impact Earth"		
Asteroids and Craters	"experimenting with different types of impacts"		
	"I loved finding and comparing craters"		
	"we got to change/use the asteroids by ourself and not be shown"		
Independence	"independent learning"		
	"we didn't have too much help – only when we needed it"		
	"we were able to experiment on the computer"		
IT Element	"working on the iPads"		
	"using iPads and a website that get us involved and learn at the same time"		
	"blowing stuff up"		
Blowing things up	"blowing up out school"		
	"blowing up everything"		
	"I learnt new things"		
Learning	"learning about asteroids"		
	"finding out what the Universe is like"		

 Table 26 - What did you like most about the activity? (Down2Earth)

Table 27 - what did you like <i>least</i> about the activity? (DownZEarth)		
Theme	Responses	
	"writing"	
Writing	"writing stuff down"	
	"writing up the results"	
	"I wish there was more choices"	
Restrictions	"being restricted on what we can and can't change"	
	"only being able to change certain things"	
	"quite repetitive"	
Repetitive	"having to repeat 9.82 x 10^{21} every time"	
	"it was very repetitive"	
	"we could have done more investigative work"	
Investigation	"we could investigate more"	
	"we didn't have to think too much"	
	"how hard it was"	
Difficulty	"it was a bit complicated at times"	
	"I didn't understand the data collecting"	
	"people messing around"	
Distractions	"people distracting me"	
	"distractions"	

Table 27 - What did you like *least* about the activity? (Down2Earth)

1 able 28 -	now do you think the activity could be improved? (Down2Earth)
Theme	Responses
	"spend a bit more time on the activity"
Time	"let us have more time"
	"more personal investigation and time"
	"more fun and interactive activities"
Additional Tasks	"maybe a bit more of a variety of practical work"
	"more simulation activities"
	"more investigation"
Investigation	"by experimenting with all aspects, not just projectile velocity"
	"it could be improved by doing more experiments"
	"more team work"
Group Work	"being able to work with a partner"
	"more group work"
	"more materials on the impact calculator"
Impact Calculator	"instead of having two planets, have them all"
	"we could have more options with regards to the projectile"
	"less written work"
Writing	"less writing"
	"less writing"
	"help us more"
Guidance	"explain the significance more"
	"be guided about what to do more"
	"more freedom"
Freedom	" be able to change everything and not restrict us"
	"give them more freedom on how to plan the investigation"
	"formulas to get the biggest crater"
Other	"learn more about astronomical objects"
	"maybe a bit more variety of practical work"

 Table 28 - How do you think the activity could be improved? (Down2Earth)

Table 29 - What did you like <i>most</i> about the activity? (Hubble)				
Theme	Response			
The Age of the Universe	"arriving at the answer of the age of the Universe using values I created"			
	"Being able to calculate the age of the Universe from real data"			
	"learning how to calculate the age of the Universe"			
	"plotting the graph"			
Graph Plotting	"the way the graph formed the results"			
	"constructing the graph"			
	"understanding real life values"			
Working with Real Data	"calculations and real life investigations"			
	"it was interactive, using real data. This makes it feel more relevant"			
	"Finding Hubble's constant"			
Finding Hubble's Constant	"calculating Hubble's constant"			
	"Learning about the Hubble constant and the Universe"			
	"the teamwork aspect of working together on an interesting topic"			
Group Work	"working as a group on a new task"			
	"working in groups"			
	"doing the calculations"			
Mathematical Aspects	"I enjoyed the mathematical section"			
	"the maths aspect"			
	"the independence"			
Independence	"the independent part – we weren't just talked through it"			
	"the freedom of finding things out on our own"			

Table 29 - What did you like *most* about the activity? (Hubble)

Table 30 - What did you like <i>least</i> about the activity? (Hubble)			
Theme	Response		
	"plotting the graph"		
Graph Plotting	"using large varieties of numbers which proved difficult to plot"		
	"finding gradient"		
	"having to repeat the same task many times"		
Repetitive	"repetitive calculations"		
	"slightly repetitive"		
	"Excel not working"		
IT Elements	"windows of laptops \mathfrak{S} "		
	"Microsoft Excel"		
	"the math at the end"		
Mathematical Aspects	"I would've enjoyed more mathematical aspects to the session"		
	"the maths"		
	"the lack of calculators available"		
Calculators	"not having a calculator"		
	"not using a calculator"		
	"the time – I wish we had longer"		
Other	"boring task"		
	"we focused on one thing, and only touched upon others"		

Table 51 - How do you think the activity could be improved: (Hubble)				
Theme	Response			
	"use computers"			
IT Elements	"less computer work"			
	"rather than using an automatic excel formula, could've created one ourselves"			
	"more time"			
Time	"needed more time"			
	"more time"			
	<i>"more interactive or rather more exploration rather than copy and paste"</i>			
Interactivity	"more practical work"			
	"more interactive"			
	"have less sheets, explain step by step"			
Instructions / Guidance	"more explanation into the equations used"			
	<i>"clearer instructions"</i>			
	"discuss more topics and add additional information"			
Other	"find out own information about our supernova"			
	<i>"it should be more expansive and include other aspects of astronomy and astrophysics"</i>			

Table 31 - How do you think the activity could be improved? (Hubble)

Appendix 15 – Students Responses to what they Liked Most about the D2E Activity relating to Context

The simulation software	(Male, Group Pink)
The website	(Male, Group Pink)
Using the simulator	(Male, Group Pink)
I enjoyed looking at how asteroids impact earth and the simulator	(Male, Group Pink)
Experimenting different types of impacts	(Male, Group Pink)
The bit with asteroids	(Male, Group Pink)
Comparing craters to the size of buildings	(Female, Group Green)
The simulator	(Male, Group Green)
Learning the properties of craters	(Male, Group Green)
Finding out that different substances create different size craters eve speed and trajectory are different	n if the size, density and (Female, Group Green)
Blowing stuff up	(Female, Group Green)
Blowing up my friends	(unknown, Group Green)
Craters	(Male, Group Green)
Blowing up our school (finding out size of craters)	(Male, Group Green)
Asteroids can be made of different materials	(Female, Group Green)
Blowing up countries	(Female, Group Green)
Seeing the crater sizes	(Male, Group Green)
Blowing up QPR	(Male, Group Green)
Creating the meteorites/asteroids and making the craters	(Male, Group Green)
Making our own crater	(Male, Group Green)
The simulations	(Male, Group Green)
Blowing up everything	(Male, Group Green)
Blowing up our school	(Male, Group Green)
I loved finding and comparing craters	(Female, Group Green)
Learning about asteroids	(Male, Group Navy)
The experiments we did and I enjoyed using the impact calculator	(Male, Group Navy)
Finding out what the universe is like	(Male, Group Navy)
Seeing how big the craters are	(Male, Group Navy)
It was based on my favourite subject	(Male, Group Navy)
Investigating and using the impact calculator	(Male, Group Navy)

Using the impact calculator	(Male, Group Navy)
It was based on space	(Male, Group Navy)
Finding out how craters hit the earth at fast speeds	(Male, Group Navy)
The impact calculator	(Male, Group Navy)
Using impact calculator	(Male, Group Navy)
Using the impact calculator	(Male, Group Navy)
Finding the size of craters	(Male, Group Blue)
The simulator	(Male, Group Blue)
Seeing how the size of the crater is effecting by the projectile velocity	(Female, Group Blue)
The simulator	(Male, Group Blue)
Seeing the size of the craters	(Male, Group Blue)
Seeing the effects of the crater	(Female, Group Brown)
Using the simulator	(Male, Group Brown)
That we could create our asteroid	(Female, Group Brown)
Just every bit as it involved astronomy	(Male, Group Brown)
Using the website	(Male, Group Brown)
We got to choose what place we blow up	(Female, Group Brown)
Seeing how much projectiles can destroy	(Male, Group Brown)