Supporting Science Foundation Year students to construct arguments from multiple documents

Caroline Anderson BSc MSc PGCE

Student ID No. 14228840

Thesis submitted to the University of Nottingham for the degree of Professional Doctorate in Education

Abstract

This research explores the experiences of students on a Science Foundation Year (Science FY) course when trying to construct arguments about controversial socio-scientific issues using multiple documents. It sought to identify the challenges these students face, design and implement an approach to support them, and explore their experiences and outcomes of this approach. Four factors from the reviewed literature, prior educational experience, disciplinary knowledge and domain-specific beliefs, justification beliefs, and topic interest informed the design principles (DPs) of the intervention. These were: *DP1* - Provide students with opportunities to develop their disciplinary knowledge and domain-specific beliefs, DP2 - Allow students to choose topics they are interested in, *DP3* – Teach students how to construct arguments, *DP4* – Equip students with the cognitive habits needed to engage with multiple documents, and *DP5* – Provide positive, timely feedback which motivates students to engage in effortful processing of belief-inconsistent information. These design principles were implemented during workshops and individual tutorials of a compulsory module of the Science FY course, through teaching, learning, assessment, and feedback activities.

A design-based research approach was adopted so that the design principles and intervention could be refined based upon student experiences. These experiences were explored using justification belief questionnaires, a piece of reflective writing, and interviews. 25 students in Study 1 participated in the intervention and six interviewees with different patterns of justification beliefs provided in depth feedback. These participants' experiences of and outcomes from the first iteration of the intervention led to the design principles and their implementation being revised. The amended activities were evaluated during Study 2 the following year with 30 students, of which seven were interviewed.

A key finding from this research is that many of the researched students had not yet developed the cognitive habits needed to manage multiple documents and did not allocate enough time, leading some to experience negative affect and adopt maladaptive strategies, such as procrastination. Another is that the abstract nature of the terms associated with components of arguments hindered understanding of

argument structure. Participants differed in their responses to designed workshop activities, but the individual tutorials were very well-received. The feedback provided during these tutorials played a vital role in motivating students to persist in the effortful processing of multiple documents. It is proposed that learning how to judge expertise, how to assess the validity of alternative perspectives, and how to use information from documents as evidence in support of claims has the potential to bring about adaptive changes to justification beliefs. These findings enable this research to make contributions to knowledge, methodology, and professional practice.

Acknowledgements

I am immensely grateful for the superb guidance, support, and encouragement I have received from those who have supervised and mentored me. I have also benefitted greatly from having numerous colleagues and peers who have taken interest in my research. Talking to them has been invaluable in developing my thinking, sharpening my analysis, and distilling my ideas.

Beyond the academic environment, completing a doctorate alongside full-time employment has only been possible because of the enduring love and unremitting support of my husband, family, and friends.

1 Table of Contents

Table of Tables

Table of Figures

Glossary of Terms

A Levels: Qualifications studied by 16- to 18-year-olds in schools and further education colleges in England, Wales, and Northern Ireland.

Beliefs in justification by authority: Students with beliefs in justification by authority seek the views of those with expertise to determine about which knowledge claims are likely to be true.

Beliefs in justification by multiple sources: Students with beliefs in justification by multiple sources look for corroborating evidence within multiple documents.

Beliefs in personal justification: Students with beliefs in personal justification believe that knowledge claims can be justified by means of personal experiences.

BTECs: Vocational qualifications which are often studied instead of, or sometimes alongside, A levels, but that have a higher proportion of coursework and practical assessment than A levels.

Department for Education (DfE): A Department of the UK Government responsible for education, training, child protection, and child services.

Foundation Year (FY): One-year course completed before an undergraduate degree by students who do not meet the requirements for direct entry onto their chosen course.

Multiple Document Comprehension (MDC): The ability to generate a coherent mental representation of source and content information relating to at least three documents.

Office for Students (OfS): Independent regulator of higher education in England.

1 Introduction

1.1 Introduction

I am a practitioner researcher. I begin this chapter by describing my professional background and explaining why I chose to research how to support students with differing justification beliefs to construct arguments from multiple documents. I then outline the four research questions I devised in relation to this research aim and introduce the methodology used to address them. The approach I took to support students was underpinned by my pedagogical philosophy, which is also expressed within this chapter. This chapter ends with a description of the structure of the thesis.

1.2 Professional Background

I am an educational professional with 25 years of experience of teaching and supporting learning. I developed my subject expertise through undergraduate and taught postgraduate study in the natural sciences and learnt how to teach these subjects through completing a Postgraduate Certificate in Education. During my 15-year career as a secondary school science teacher, I taught Biology, Chemistry, and some aspects of Physics to children aged 11 to 16 years old, and Biology and Psychology to learners aged 16 to 19.

During the past 10 years, I have been teaching and supporting students who are studying biological and biomedical sciences at university. When I first joined the university where this research was conducted, my role was to support Year 1 and 2 undergraduate science students in completing written assessments. I noticed that many of these students found working with multiple documents to be very challenging, particularly if those documents featured differing perspectives or provided conflicting information. Even when documents were concordant, students struggled to synthesise the sources. This often resulted in sources being treated as independent entities within students' essays. As such, their work did not communicate how the findings of the various studies had collectively contributed to our knowledge of a particular aspect of science.

My role at the university changed five years ago and I became the Co-Course Director for the Science with Foundation Year (Science FY) course. This is a year-long course for students who are not eligible for direct entry onto undergraduate science degrees. I observed that Science FY students often lack awareness of what constitutes evidence in science and often select inappropriate sources, such as personal accounts, when asked to provide evidence in support of claims. Having begun to read the literature about multiple document comprehension and argumentation to try to better understand how to support students in reading multiple documents and constructing arguments, I soon encountered research about epistemic cognition. I was particularly drawn to research about how students justify what they believe to be true and the impact of different justification beliefs on engagement in argumentation and multiple document comprehension. Consequently, I decided to focus my research on designing teaching, learning, assessment, and feedback activities which would support students with differing justification beliefs in constructing arguments from multiple documents.

Throughout my career in education, I have sought opportunities to develop adaptive expertise (Hatano & Inagaki, 1984). This is an attribute of lifelong learners that enables them to apply their knowledge and skills in different situations. Completing a Professional Doctorate in Education has given me another opportunity to expand the breadth and depth of my expertise. One of the most significant challenges I have faced while completing my research has been to reconcile my training as a natural scientist, where the goal was to seek statistical generalisation (Bakker & van Eerde, 2015), with my experiences as an educator where I have observed how differently individuals respond to the same educational experience. I decided to adopt a mixed methods approach, where I was able to draw upon my existing knowledge of quantitative data analysis while also learning how to generate and analyse qualitative data.

Attaining my Professional Doctorate in Education will be a significant personal achievement and a very proud moment for my family as I will become the first to attain at this level. It will also cement my professional status as a highly experienced educator whose practice is informed by research as well as experience.

1.3 Research Aim, Questions, and Methodology

The overall aim of this research was to design, implement, and explore students' experiences of and outcomes from an educational intervention that sought to support Science FY students with differing justification beliefs in constructing an argument from multiple documents.

Four research questions were devised in relation to this aim:

RQ1: What challenges do Science FY students face when constructing arguments from multiple documents?

RQ2: What are the justification beliefs of these students?

RQ3: What are the key design principles of an intervention designed to support these students in constructing an argument from multiple documents?

RQ4: What are students' experiences of and outcomes from this intervention?

These research questions were addressed using the methodology of design-based research. Design-based research requires iterative phases of data generation and analysis. This research used a questionnaire to measure participants' justification beliefs and to provide one means of evaluating the outcome of the intervention. Data generated from audio recordings of individual tutorials, interviews, and a piece of reflective writing enabled participants' experiences to be explored. These data also aided the interpretation of the changes to participants' justification beliefs.

1.4 Pedagogical Philosophy

My approach to teaching and supporting learning is oriented towards learners' future potential. My desire to enable learners to master complex tasks they are not yet capable of achieving independently is founded in scaffolding (Wood et al., 1976), which is central to my pedagogical philosophy. As such, the implementation of the educational intervention described within this thesis was informed by scaffolding.

I first encountered the writings of Lev Vygotsky (1896-1934) when studying developmental psychology during my undergraduate degree studies. Although now considered by some (for example, Newman & Latifi, 2021) to be problematic, during my teacher training, Vygotsky's various writings about the zone of proximal development were related to scaffolding, a concept that originated from Wood et al.'s (1976) observations of young children (3–5-year-olds) interacting with an adult. Interestingly, part way through my doctoral studies, I discovered that David Wood had supervised my primary supervisor, Prof Shaaron Ainsworth.

Wood et al.'s (1976) observations led them to describe how during interactive, instructional relationships, the adult, or tutor takes charge of those aspects of a complex task that are initially beyond the child's, or learner's, ability. The tutor has a theory of the task; that is, they understand how it can be completed. The tutor also has a theory of the learner, which allows them to understand why the learner may not always follow the guidance given. Together, the theory of the task and the theory of the learner enable the tutor to diagnose the learner's current capabilities and provide appropriate feedback and guidance that enables the learner to develop further capabilities. The tutor hands over responsibilities to the learner once they can complete the task unaided.

The strategies used by the tutor to support the learner are both motivational and cognitive (Belland et al., 2013). It has long been known that motivated students are more likely to be emotionally, cognitively, and behaviourally engaged in their studies (Ryan & Deci, 2000). Acosta-Gonzaga and Ramirez-Arellano (2022) recently used structural equation modelling to explore the relationships between scaffolding, motivation, emotional engagement, cognitive engagement, and behavioural engagement in German university students. Their findings indicate that scaffolding promotes the adoption of more elaborative learning strategies and enhances emotional engagement. Students who experience positive affect are more behaviourally engaged and are more likely to have higher learning achievements. Thus, scaffolding is likely to have positive effects on students' cognitive habits, their affective state, their engagement in their studies, and their levels of attainment.

Although scaffolding was originally conceived as a means of describing dyadic interactions, like many educational practitioners I typically employ scaffolding in whole-class situations. Smit et al. (2013) identified three features of whole class scaffolding during their observations of the interactions between Dutch primary school teachers and the children in their classes. Firstly, scaffolding is layered; that is, teachers diagnose learners' capabilities, give feedback and guidance, and conceive ways to foster independence through their actions in and out of the classroom. Secondly, scaffolding is distributed over several lessons. Thirdly, scaffolding in whole class settings is cumulative, with repeated diagnoses of learners' capabilities and the provision of feedback and guidance sustaining the process of handover to independence. Even though I have always taught learners older than those in Smit et al.'s study, much of what they observed is also present in my practice. I pay attention to learners' initial capabilities both inside and outside of the classroom. In response, I amend my teaching and develop learning resources to better match what learners' can and cannot yet do. I also adapt how I interact with learners in the classroom, providing those at an earlier stage of skill acquisition with more direct verbal feedback and guidance than those who are close to mastering the task, with the aim of enabling all learners to be able to complete complex tasks independently. These characteristics of my practice are evident within the descriptions of the workshops, individual tutorial, and assessment activities that constituted the educational intervention.

1.5 Structure of the Thesis

This thesis has seven further chapters. Chapter 2, Educational Context, begins with an overview of Foundation Years in UK Higher Education before describing the Science with Foundation Year programme at the researched university. This chapter also gives details of the compulsory module within which the educational intervention described within this thesis was embedded. Attention is drawn to changes made to this module as a result of this research.

Chapter 3 is the Literature Review. Its purpose is to introduce the key terms and influential models associated with research about argumentation, multiple document comprehension, and epistemic cognition. This chapter also considers

why science students need to be able to engage in argumentation, offering justifications at the individual, community, and societal levels. The reviewed literature surfaced four factors thought to contribute to the challenges students face when constructing arguments from multiple documents: prior educational experience; disciplinary knowledge and domain-specific beliefs; justification beliefs; topic interest. These factors informed the development of five design principles: *DP1* - Provide students with opportunities to develop their disciplinary knowledge and domain-specific beliefs; DP2 - Allow students to choose topics they are interested in; *DP3* – Teach students how to construct arguments; *DP4* – Equip students with the cognitive habits needed to engage with multiple documents; *DP5* – Provide positive, timely feedback which motivates students to engage in effortful processing of belief-inconsistent information.

Chapter 4 gives details of the methodology and methods used to generate data. It begins by describing the pragmatic aim of this research and explains why the pragmatic approach aligns with the methodology of design-based research. This chapter also describes the quantitative and qualitative methods of data generation used within this research, including providing details of the materials used, and how these were embedded into the different stages of the design-based research process. There are also details of the statistical techniques used to analyse the quantitative data and of the Framework Analytical Process (Ritchie & Spencer, 1994), which enabled the creation of a thematic framework from the qualitative data. This thematic framework was used to compare the experiences of those with similar and different justification beliefs. Chapter 4 concludes with consideration of the ethical issues faced when the researcher is also the practitioner and explains how ethicality enabled me to juggle the dual roles I played within this research.

Chapter 5 returns to the design principles first outlined in Chapter 3, though here the focus is on how these design principles influenced the design of the educational intervention. The first iteration of this intervention featured three workshops, an individual tutorial, and reflection and action planning activities. Chapter 5 provides details of the teaching, learning, and formative feedback activities that took place within the workshops and individual tutorial. It also

explains how these activities, along with the post-tutorial reflection and action-planning activities, aimed to address the design principles (*RQ3*).

Chapters 6 and 7 are primarily concerned with the findings from the two iterations (Study 1 and Study 2) of the intervention. Both chapters begin by describing participants at the cohort level before providing detailed descriptions of the justification beliefs of those selected as interviewees (*RQ2*). Statistical analysis of the changes to justification beliefs provides one of the outcome measures in relation to *RQ4*. Study 1, which is described in Chapter 6, was primarily concerned with the challenges faced by students when trying to construct arguments from multiple documents (*RQ1*); the thematic framework developed within this research to discuss these challenges is explained. The Study 1 interviewees were also asked to describe their experiences of the intervention activities. Their responses informed changes to the design principles and their implementation in Study 2. These changes are described at the start of Chapter 7. Since the purpose of Study 2 was to explore students' experiences of and outcomes from the intervention, much of Chapter 7 reports the findings in relation to *RQ4*.

In Chapter 8, I summarise the extent to which the research has achieved its aims by discussing the key findings in relation to the four research questions, consider how these findings relate to the literature, and indicate the contributions made to knowledge, methodology, and professional practice by this research. Chapter 8 also draws attention to the limitations of this research, provides recommendations for future research and considers the implications of this research for future practice.

2 Educational Context

2.1 Introduction

The research described within this thesis was conducted in the university where I am employed as an academic member of staff with responsibilities for teaching and curriculum leadership of the Science Foundation Year (Science FY) course. This chapter begins by providing a broad overview of Foundation Years (FYs) within UK higher education. It then becomes focused on the Science FY course at the researched university. This course is comprised of four compulsory year-long modules. Three of these modules are subject-focused. The fourth, Studying Science at University (Studying Science), is skills-based; it is within this module that the educational intervention described within this thesis is embedded. This chapter describes the aims of this module and provides details of its assessment activities.

2.2 Foundation Years in UK Higher Education

In the UK, FYs are one-year courses completed before an undergraduate degree by students who do not meet the requirements for direct entry onto their chosen course. Many FYs are integrated into undergraduate degree programmes, but some are offered as standalone programmes. Data reported in 2019 by the Office for Students showed that the number of entrants to integrated FYs in 2017/18 was 30,030, of which just under a fifth were studying biological or biomedical subjects. UK students wishing to study a FY usually apply through the Universities and Colleges Admissions System (UCAS). A course search on the UCAS website for 2018 entry indicated that there were 79 FY courses enabling progression to degrees in the biological, biomedical, and natural sciences. These courses were being offered at 29 UK universities.

2.3 Science with Foundation Year Course

The teaching year at the university where this research took place is organised into semesters and terms. Semester 1 begins in late September and ends in late January; Semester 2 begins in late January or early February and ends in late June. The terms describe the periods within the semesters when teaching and assessment

activities take place. The autumn term begins in late September and ends in mid to late December. The spring term begins after the winter break in mid-January and ends in mid to late March, or early April; the timing depends on Easter. The summer term begins after the spring break in mid to late April or early May and ends in late June.

The Science FY is comprised of four compulsory modules that span across both semesters. Two of these modules, Foundation Biological Sciences (Biology) and Foundation Chemistry (Chemistry), are designed to cover the core content of A Level Biology and A Level Chemistry programmes of study, respectively. A Levels are qualifications studied by 16- to 18-year-olds in schools and further education colleges. The UK Government's Department for Education mandates the knowledge, understanding, and skills that must be addressed within an A Level programme of study. This core content accounts for 60% of an A level programme. A Levels typically require two years of study and are assessed through coursework, practical assessments, and examinations. They are graded from A* to E, with A* being the highest.

The Maths for Foundation Science (Maths) module has a different aim to that of the Biology and Chemistry modules; it seeks to develop students' confidence in applying mathematical principles to scientific applications. The fourth module, Studying Science, is skills-based, with students being assessed on their levels of competence rather than their ability to recall and apply subject content. Further details of the Studying Science module are given in the next section as it was within this module that the research took place.

2.3.1 Science FY Students

A typical cohort of the Science FY at the researched university has between 30 and 45 students. Many have studied one A level science subject, often Biology, but lack the second science subject typically needed for entry onto degrees in the biological, biomedical, or natural sciences. Others have studied A levels in subjects other than the sciences. To gain entry onto the Science FY course, students must have achieved grades ABC, BBB, or BBC in their A levels. Other grade combinations, such as A*BD or A*CC, are usually accepted though these are considered on a case-by-

case basis. Students awarded grade E in any of their A levels are only admitted in exceptional circumstances.

Some students on the Science FY have not studied A levels but have instead achieved vocational or technical qualifications at an equivalent standard, which is known as Level 3. Many schools and colleges offer vocational or technical qualifications alongside A levels, though some further education colleges specialise in these types of qualifications. Of these vocational or technical qualifications, Level 3 BTEC Nationals (BTECs) are the most common. These are two-year qualifications which consist of core and optional modules, with schools and colleges free to choose which combination of optional modules are offered to their students. BTECs have a higher proportion of coursework and practical assessment and a lower proportion of end of course examinations than A levels and as such, many undergraduate science courses at the researched university do not accept BTECs for direct entry. Students with BTECs are therefore often required to complete the Science FY prior to starting their undergraduate degree. BTECs are awarded at Distinction*, Distinction, Merit, and Pass level; those with grades at Pass level are not accepted onto the Science FY.

The Science FY also admits students who have completed Access to Higher Education courses; these are one-year courses provided by further education colleges to enable people whose qualifications do not meet the typical entry requirements for undergraduate study to gain access to higher education. Most students who complete Access courses have spent a few years outside of formal education and are described as being mature students.

During the time the research was conducted, the Science FY tended to have a relatively high proportion of non-UK students. Some of these students were from countries in mainland Europe, such as Austria, Bulgaria, Cyprus, Germany, Italy, The Netherlands, Poland, and Portugal. Others were from further afield, such as China, Ghana, Hong Kong, India, Macau, Nigeria, Oman, Saudi Arabia, Syria, Taiwan, Thailand, Turkmenistan, and the United States of America. Non-UK nationals have usually gained qualifications within their home countries, though a few have completed international A Level programmes or the International Baccalaureate (IB) Diploma programme. The IB Diploma is multi-disciplinary, with students completing

courses in languages and literature, humanities, sciences, mathematics, and the arts during their two years of study. Students take a minimum of three and a maximum of four of these subjects at higher level and the others at standard level. Each course is assessed through written examinations, externally assessed coursework, and in-school assessment tasks; the resulting grades range from 7 to 1, with 7 being the highest. These grades are summed to determine the final diploma score, with students needing to achieve a minimum of 24 points to be awarded the diploma. The Science FY course typically accepts students who have achieved a minimum of 28 points overall, with at least grade 5 in science subjects studied at standard level and grade 4 in science subjects studied at higher level.

A typical cohort of Science FY students therefore includes students who have recently completed A levels, Level 3 BTEC Nationals or equivalent international qualifications alongside mature students who have Access to HE qualifications.

2.4 Studying Science Module

The Studying Science module aims to prepare students to study effectively on a science undergraduate degree programme by developing skills and qualities that will support their transition to university and enable them to become a successful science student. These skills and qualities include making enquiries using the scientific method, developing strategies for reading effectively, making notes effectively and writing scientifically, finding, evaluating, and summarising multiple sources, and time-management.

The teaching in the Studying Science module is organised to align with the semesters. Semester 1 is primarily concerned with scientific research methods, with students learning how to formulate scientific questions, hypotheses, and predictions. An overview of the teaching, learning, and assessment activities during Semester 1 is provided shortly. In the academic year prior to this research taking place, Semester 2 was concerned with scientific communication. One aspect of this was how the mainstream media report scientific discoveries. Another was that scientists may hold different perspectives about the same issue. Through conducting this research, the focus of Semester 2 shifted to being about how scientists use evidence from scientific research to support claims made when

engaging in argumentation. The section about Semester 2 therefore firstly describes how this part of the module originally operated and then explains how the module changed when the designed intervention was implemented.

2.4.1 Semester 1

The aim of Semester 1 is for students to develop their understanding of how scientists study the world around them. They are taught about how scientists devise hypotheses and make predictions based on existing scientific theories, models, and laws. They learn about different experimental designs and data collection methods and are introduced to the ethical guidelines that aim to reduce the risk of causing harm to participants or researchers. They gain experience of presenting and interpreting scientific images and data. They also learn about sources of error and the importance of controls in scientific experiments.

At the time the research was conducted, students were assessed towards the end of Semester 1 through a timed assessment completed under examination conditions. The examination had two sections. Section A featured three or four unseen short answer questions. These questions assessed students' knowledge of key definitions, their knowledge and understanding of key concepts associated with the scientific method, including ethical considerations, and their ability to interpret and evaluate data obtained using scientific research methods. Section B was a 'seen' extended answer question, which means that students were given access to this question in advance of the examination so they could prepare their response to it. Specifically, they were asked to formulate a research question, outline their hypothesis, identify the variables, and state their prediction. They were then asked to provide a detailed description of how they would test their prediction. The prompts indicated that their description should identify and justify their choice of data collection method, and whether data would be collected in the lab, in the field or in silico. They were also guided to provide details of the sample or participants and explain how the participants or researcher, as appropriate, would be kept safe from harm with reference to appropriate guidelines. In addition, they were prompted to describe the procedure and identify aspects of this procedure that would enable the collection of valid and reliable data.

2.4.2 Semester 2

In the academic year prior to this research taking place, the main purpose of Semester 2 was to consider how advances in scientific understanding are communicated to the public. The Semester 2 assessment required students to select a media story about a recent scientific discovery and consider whether the claims made within the story were supported by evidence obtained using scientific research methods. Students were able to select their own media story, but there were two constraints. Firstly, the scientific discovery featured in the story had to relate to the degree subject they wished to progress onto. For example, students wishing to study Pharmacy had to select a story relating to preventing or treating illness. Secondly, there needed to be different perspectives about the discovery. For example, the media story may indicate that while some were enthusiastic about the potential of a new treatment, others were more circumspect.

As indicated in Table 2-1, Semester 2 featured three 'Science in the media' workshops to engage students with examples of how scientific information is communicated to the public. The first of these workshops, 'Science in the media: Understanding data' required students to consider how data from scientific research are presented in the media, whether these representations match the conventions within the scientific community, and critically analyse whether the data are represented accurately. To aid students in interpreting data, they were also taught fundamentals of descriptive and inferential statistics during this workshop.

The second workshop, 'Science in the media: Case 1 – The MMR controversy', introduced students to the idea that the behaviour of the public can be influenced by media reporting of scientific discoveries. As indicated by its name, the controversy surrounding the combined measles, mumps, and rubella (MMR) vaccine and whether this vaccine increased the risk of children developing autism featured prominently within this workshop. This case was, at the time, being discussed in the mainstream media because the number of cases of mumps had been rising since 2005, and there was debate about whether this increase was occurring because a lower proportion of children born in the early 2000s had been fully vaccinated against measles, mumps, and rubella. It should be noted that

opinions about the safety of mass vaccination programmes were not as divergent as they became during the Covid-19 pandemic, nor were they expressed as strongly.

The third Science in the media workshop, subtitled 'Case 2 – The 'wonder' drug' continued to explore how media reporting of scientific discoveries influences the public, with students being asked to consider whether claims made in the media about so-called 'wonder drugs' are justified. They were also directed to consider how such discoveries are reported and whether media reporting contributes positively to public perceptions of science. The example used within this workshop was that of a drug used to treat a rare genetic disorder which causes premature ageing. Students considered whether the findings of a study about this drug supported a claim made within a popular lifestyle publication that this drug could prolong the lifespan of those without the condition.

Students were assessed by means of a group poster and a piece of reflective writing. To help students develop their group working skills, the university's Careers and Employability Service ran a 'Working in teams' workshop in Week 1 of Semester 2. This was followed a few weeks later by a 'Poster groups planning session', during which students were reminded of the key messages from the 'Working in teams' workshop. Each group was subsequently invited to two tutorials to discuss their progress in creating the group poster. Students were expected to complete an individual 'Reflecting on Feedback' assignment after the first of their poster group tutorials.

Semester 2 also featured two workshops to support students in finding the sources needed to complete their group poster. The first of these, 'Selecting library resources for your group project', was delivered by the university's library team and introduced students to online search tools such as Google Scholar and subject-specific databases such as Web of Science. The second, 'Effective reading strategies', provided an overview of the structure of journal articles and where to find key pieces of information within them. There was also a workshop about 'Managing your time effectively', during which students were encouraged to map their coursework deadlines onto a termly planner.

Table 2-1 *Teaching, Learning, Assessment, and Feedback Activities During Semester 2 of the Academic Year Prior to the Research Being Conducted*

Teaching week in Semester 2	Teaching, learning and feedback activities		Assessment activities
1	Introduction to the spring semester assessments	Working in teams	Group poster assignment set
2	Managing your time effectively	Science in the media: Understanding data	
3	Selecting library resources for your group project	Science in the media: Case 1 – The MMR controversy	
4	Effective reading strategies	Science in the media: Case 2 – The 'wonder' drug	
5	Poster groups planning session		
6	Poster group tutorials		
7	Poster group tutorials		
8	Guidance about the 'Reflecting on Feedback' assignment		Individual 'Reflecting on Feedback' assignment set
9	Poster group tutorials		
10	Poster group tutorials		'Reflecting on Feedback' assignment submitted
11			
12			Group poster presentation took place

The research described within this thesis led to substantial changes being made to the activities and assessments in Semester 2. Most notably, the focus shifted from being about the role the media play in communicating scientific ideas to being

about how scientists use evidence from scientific research to support claims when engaging in argumentation. The poster activity about how a scientific discovery was presented in the media was removed; students instead complete the 'Making an Argument' assignment. This individual assignment requires students to select a socio-scientific issue, identify the different perspectives about this issue, and construct an argument in support of one of these perspectives. Sadler and Ziedler (2005) describe socio-scientific issues (SSIs) as being controversial issues which provoke social debate about the processes or products of science. SSIs are often subject to political and ethical influences. Engaging with SSIs is thought to increase student motivation, retain student interest, and develop students' higher order thinking skills (Sadler, 2009). Students are guided to select an SSI that relates their career aspirations, personal interests, or academic interests. The SSI must have featured in the media in some form during the last twelve months. Table 2-2 shows a selection of the SSIs chosen by Science FY students in recent years. By striving to engage Science FY students with SSIs, Semester 2 of the Studying Science module differs from more traditional models of science education that see scientific knowledge as something created by experts and transmitted by teachers to learners (Duschl & Osborne, 2002). Instead, students are challenged to develop their scientific understanding through synthesising information from multiple documents and constructing an argument about a contemporary issue.

Table 2-2 *Examples of Socio-Scientific Issues Selected by Science FY Students, Organised into Five Broad Themes*

Theme	Example socio-scientific issues		
Health	Anti-cholinergic drugs and their links with dementia	The use of oestrogen-progesterone preparations as hormone replacement therapy	Mice as model organisms for studying anxiety
Nutrition and diet	The effectiveness of intermittent fasting as a means of weight loss	The effectiveness of the ketogenic diet in managing chronic health conditions	The safety of artificial sweeteners
Sport	The effectiveness of myofascial release and static stretching in improving muscle performance	The use of performance enhancing drugs in sport	Fair competition between birth-gendered and trans-gendered female athletes
Agriculture and environment	Changes to farming practices to combat climate change	The potential of lab grown meat to provide a sustainable food source for the global population	The case for rewilding in the United Kingdom
Ethical dilemmas	Alternatives to animal testing for scientific and medical purposes	The case for an international moratorium on germline gene editing	The use of embryonic cells for stem cell research

Table 2-3 provides an outline of Semester 2, showing how the three workshops and individual tutorials which feature the intervention activities designed through this research have been scheduled alongside three sessions retained from the previous academic year.

Table 2-3 *Revised Teaching, Learning, Assessment, and Feedback Activities During Semester 2*

Teaching week in Semester 2	Designed intervention activities	Activities retained from previous academic year
1	Workshop 1: Controversial socio-scientific issues	Introduction to the spring semester assessments
2	Workshop 2: Constructing arguments	
3	Workshop 3 Reading and evaluating scientific papers	
4		Selecting library resources
5	Individual tutorials	Guidance about the Reflective Commentary assignment
6	Individual tutorials	
7	Individual tutorials	

The first session in Semester 2 introduces the Making an Argument and Reflective Commentary assignments. The introductory session is followed by three workshops that feature designed intervention activities. The Selecting library sources, which takes place after these workshops and is delivered by the university's library staff, was retained from the previous academic year because it introduces students to search engines and databases. The workshop giving guidance about reflecting on feedback has also been retained. This workshop continues to describe the Gibbs (1988) structured debriefing framework and explain how this framework aligns with the three questions students answer when completing what is now known as the Reflective Commentary assignment. This assignment requires students to reflect on the individual feedback they are given about their Making an Argument assignment (for further details, see Chapter 4).

2.5 Summary

The educational intervention described within this thesis is embedded within Semester 2 the Studying Science module of the Science FY. The development of this intervention brought about substantial changes to Semester 2, which shifted from

being about how scientific discoveries are communicated to the public to become focused on how scientists use evidence to support their claims when engaging in scientific arguments. New workshops, tutorials, and assessment activities were designed to support students in constructing an argument from multiple documents.

3 Literature Review

3.1 Introduction

Students studying science at university can find constructing arguments to be very challenging. This is partly because many students lack prior experience in reading and writing scientific arguments. It is also because reading, understanding and synthesizing information from multiple documents is difficult. Additionally, the ways in which students justify what they believe to be true can be maladaptive. The research described within this thesis draws upon three areas of literature, namely, argumentation, multiple document comprehension, and epistemic cognition, to identify the difficulties students face when constructing arguments and to suggest ways in which they can be supported.

The intersections between these three areas of literature are of particular importance. The models of multiple document comprehension and epistemic cognition which have influenced this research operate within these intersections. These models are Richter and Maier's (2017) two-step validation model, List and Alexander's (2017) Cognitive Affective Engagement Model, Kuhn's (1991) model of informal reasoning, Greene et al.'s (2008) Epistemic and Ontological Cognitive Development Model, and Ferguson et al.'s (2012) Justification for Knowing model.

Engaging in argumentation is believed to beneficial for science students, their participation in scientific communities, and wider society. It is therefore concerning that many students lack prior educational experience of engaging in argumentation. Students may also face other barriers, such lacking prior knowledge of the discipline, having maladaptive justification beliefs or being disinterested in the topic. Students can, however, be supported in constructing arguments from multiple documents. The final section of this literature review introduces five key design principles that were identified from the literature and have informed the design of the intervention.

3.2 Argumentation, Multiple Document Comprehension, and Epistemic Cognition

Three key areas of literature underpin this thesis. This section defines key terms associated with each area of literature and considers the most influential models within each of these areas. Many of these models illustrate how the different areas of research interact with each other, thus attention is drawn to the intersections between argumentation, multiple document comprehension and epistemic cognition.

3.2.1 Argumentation

Duschl and Osborne (2002) define argumentation as the process of constructing an argument. Arguments are often constructed using rhetorical frameworks. These frameworks typically consist of a claim and an accompanying structure that attempts to convince the reader to accept the claim (Venville & Dawson, 2010). This research adopts the most widely used of these frameworks: Toulmin's (1958) model of practical arguments. Practical arguments typically begin with the claim; this is a conclusion, proposition or assertion which may be controversial but can be justified on the basis of what is known as the ground, datum or, more commonly, evidence. The relationship between the evidence and the claim may not always be obvious, hence arguments often require warrants that explain how and why the evidence supports the claim. The inclusion of a warrant is particularly important in scientific writing to explain why a particular claim can be justified from the evidence (von der Mühlen et al., 2019). The warrant may itself be underpinned by assumptions or supported by evidence; these aspects of an argument are known as backings. Even though the aim of an argument is to convince the reader to believe the claim, it is important to indicate the degree of certainty in the claim by adding qualifiers such as modal verbs (for example, may) and uncertainty markers (for example, likely) where needed. There may also be situations in which restrictions are applied to the claim; rebuttals are used to indicate such circumstances.

Toulmin's (1958) framework is not without its critics. For instance, Andriessen and Baker (2014) rightly note that the model is primarily concerned with the relationships between pieces of knowledge and largely ignores the personal,

socio-relational, and emotional aspects of argumentation. Toulmin's (1958) model and its accompanying examples have also been criticised for being largely abstract, with the evidence featuring within the accompanying examples differing somewhat from what the scientific community would regard as evidence in science (Duschl & Osborne, 2002). Notwithstanding these valid concerns, Toulmin's (1958) model remains dominant and its use in science is appropriate, not least because scientists often use a rhetorical framework to relate new data to existing theory when attempting to convince others to accept new scientific discoveries (Duschl & Osborne, 2002).

3.2.2 Multiple Document Comprehension

The rapid rise in knowledge production and its dissemination through modern technologies make it much less likely that students will locate all the information they need to construct an argument within one document (Stadtler, 2017). Students will therefore need to access multiple documents, whose heterogeneity requires them to cope with diverse information sources that are available in different formats, accessed in a variety of ways, use different vocabulary and grammatical structures, and vary considerably in quality (Rouet & Britt, 2014; Strømsø & Kammerer, 2016). The information within these documents may contradict students' prior beliefs (Stadtler, 2017) and provide divergent explanations about situations and phenomena (Rouet & Britt, 2014). It is especially likely that science students will encounter conflicting explanations because the discursive nature of knowledge production in science allows scientists to hold differing perspectives (Stadtler, 2017).

Theoretical models of multiple document comprehension (MDC) typically assume that students will encounter conflicting information when they read multiple documents (Strømsø, 2017). The research described within this thesis draws on two models of MDC: Richter and Maier's (2017) two-step validation model and List and Alexander's (2017) Cognitive Affective Engagement Model (CAEM). Both are theoretically grounded in the Documents Model Framework (Britt et al., 1999; Perfetti, et al., 1999). It is therefore important to firstly understand how the Documents Model Framework (DMF) explains the processes that skilled readers

undertake when reading multiple documents. The DMF has two interconnected models: the integrated mental model (Britt & Rouet, 2012), and intertext model (Le Bigot & Rouet, 2007). The integrated mental model is concerned with how readers interpret the situation or phenomenon described in the document in relation to their prior knowledge and to other documents. Students who can form an integrated mental model are able to organise ideas that are found only in one document, ideas that are shared across multiple documents, and ideas that are contested in different documents. The intertext model is concerned with how students process source information, such as the author or organisation responsible for the document, the date of publication, and the author's rhetorical goal (Le Bigot & Rouet, 2007). Students who can form an intertext model are therefore able to describe who said what, when they said it, and understand why they said it. Once students have formed the integrated mental model and the intertext model, they can use rhetorical predicates such as according to, corroborates, and contradicts, to tag ideas within the integrated mental model to source information within the intertext model, thus forming a documents model.

When constructing a documents model, students often use what Wineburg (1991) described as being the sourcing and corroboration heuristics but are now more commonly referred to as sourcing and corroboration behaviours. Sourcing behaviours include attending to and evaluating the source of the document prior to reading it (Rouet & Britt, 2014). Corroboration behaviours include identifying differing perspectives and comparing claims made by one source against prior knowledge and against claims made in other documents to determine whether the claims agree or are divergent.

Le Bigot and Rouet (2007) stated that directing students to pay attention to sources enables them to create the intertext model (author, date, and rhetorical goal). While there is evidence that students spontaneously pay attention to source information, they may over-rely on their assessment of the credibility of the author when determining whether the information provided within the document is likely to be true. For example, Mason et al. (2010) found that although most students identified the author's identity and their institutional affiliation, less than half assessed whether the claims made by the author were supported by evidence.

Other research has shown that sourcing behaviours are affected by whether students are familiar with the topic they are researching. For instance, Mason and Boldrin (2008) observed that university students searching for information on the internet about an unfamiliar topic often favoured sources that matched their pre-existing views. These findings are pertinent because the researched students are expected to have some awareness of their topic but are unlikely to be highly knowledgeable.

Since students are likely to access multiple documents which have been written by different authors, the responsibility for spotting discrepant information lies with the students (Strømsø & Kammerer, 2016). Accessing information online also makes it even more likely that students will encounter deliberately misleading information (Stadtler, 2017). The documents model (Britt et al., 1999; Perfetti et al., 1999) predicts that only those students who possess effective corroboration behaviours will identify, compare, and verify conflicting information.

Although strongly influenced by the documents model, the two models of MDC adopted within this research contend that possessing sourcing and corroboration behaviours is not sufficient. Both predict that students also need to be affectively engaged, though they differ in the extent to which affect is thought to impact upon multiple document comprehension. According to Richter and Maier's (2017) two-step validation model, students who possess the cognitive capabilities to read and process texts will attend to belief-consistent information regardless of their motivational state. However, once they detect inconsistencies between new information and their prior knowledge, only motivated students will engage in the resource-intensive elaborative processing of belief-inconsistent information. Students who lack the desire to construct justified and defensible arguments, and those who do not see the need to defend their position against alternative views, are unlikely to explore alternative perspectives (Richter & Maier, 2017).

Affective engagement plays a more prominent role in List and Alexander's (2017) Cognitive Affective Engagement Model (CAEM), where it is thought to interact with cognition in all aspects of multiple document comprehension. This model predicts that students who lack affective engagement will tend to employ simple heuristics

in a perfunctory manner even if they possess the cognitive habits needed to take a more nuanced approach. For example, they may reject all websites as being untrustworthy, rather than considering the author's affiliations and expertise. Furthermore, they are likely to enact simple rules, such as seeking verification across two documents, when deciding how many documents to read. The CAEM also predicts that having high levels of affective engagement that are not matched by effective cognitive habits can also be problematic, particularly as such students are often highly motivated to read. In seeking to satiate their interest, students who lack effective corroboration behaviours risk becoming overwhelmed. It is only when students have high levels of affective engagement coupled with effective sourcing and corroboration behaviours that they will be able to successfully attend to, evaluate, and integrate source and content information from multiple documents. This is because these so-called 'critical analytic' students are motivated to access multiple documents with the aim of creating a coherent mental representation of the information found in multiple documents.

Although the two-step validation model and the CAEM differ in how they view affective engagement, both indicate that it is only when students are motivated and interested in what they are reading that they will be able to take full advantage of having effective sourcing and corroboration strategies. Hence, within this research, fostering motivation and affective engagement is viewed as being as important as enabling students to develop effective cognitive habits.

3.2.3 Epistemic Cognition

It is widely thought that epistemic cognition, which is how people think as they "acquire, understand, justify, change, and use knowledge" (Greene et al., 2016, p.1), influences learning and reasoning. It is known that epistemic cognition influences reading, thinking, and processing when reading texts (Strømsø & Kammerer, 2016). Epistemic cognition also predicts levels of criticality (Nussbaum et al., 2008), scientific reasoning (Hotulainen & Telivuo, 2015), and the ability to detect fallacies and conflicting information (Iordanou et al., 2016).

The term epistemic cognition entered the literature relatively recently, with researchers previously referring to personal epistemologies (Hofer & Pintrich,

1997), epistemological beliefs (Schommer, 1990; Schommer-Aikins, 2004), practical epistemologies (Sandoval, 2005), or epistemic beliefs (Muis & Franco, 2009). The term epistemic beliefs will be used within this thesis when specifically considering the beliefs which students hold about the nature of knowledge and the process of knowing (Hofer & Pintrich, 1997). Of particular interest is how students justify that their beliefs are true (Greene et al., 2008); hence the term justification beliefs features prominently within this thesis.

Educational research has long been concerned with epistemic cognition and many different models have been proposed. This research has been influenced by three models: Kuhn's (1991) highly influential model of informal reasoning, Greene et al.'s (2008) Epistemic and Ontological Cognitive Development Model, and Ferguson et al.'s (2012) Justification for Knowing model.

Kuhn's (1991) developmental model of informal reasoning first forged the connection between epistemic beliefs and argumentation. This model proposed that students can be absolutists, multiplists or evaluativists. Absolutists believe knowledge to be absolute and certain. Multiplists recognise that different viewpoints exist but believe all personal views to be equally valid and fail to evaluate claims. It is only when individuals develop evaluativist beliefs that they understand that viewpoints can be compared and evaluated to assess their validity. Somewhat unusually for a developmental model, Kuhn offered suggestions about how students can move from one developmental stage to the next. For example, absolutists are likely to become multiplists after interacting with others who hold differing perspectives (Hofer & Pintrich, 1997).

This research adopts the term multiplist from Kuhn's (1991) model. It also adopts terms from the Epistemic and Ontological Cognitive Development model, which was proposed by Greene et al. (2008). This model assumes that students entering university no longer believe that knowledge is simple and certain. Using concepts derived from Chandler et al.'s (1990) work on the dogmatism-scepticism axis, Greene et al. postulated that individuals respond in one of two ways when faced with the epistemic challenge of reconciling complex and uncertain information. Some seek expert's views about which knowledge claims are likely to be true. These students are thus described as dogmatists. Others disregard expert

views and instead opt to believe that knowledge claims can be justified by means of personal experiences; these students are termed sceptics. Greene et al.'s model therefore uses two dimensions of epistemic cognition, justification by authority and personal justification, to distinguish between dogmatists and sceptics. Dogmatists have strong beliefs in justification by authority and weak beliefs in personal justification whereas sceptics have weak beliefs in justification by authority and strong beliefs in personal justification.

According to Greene et al.'s (2008) model, the final stage of epistemological and ontological development sees students accepting that both personal justification and authority play a role in justifying knowledge claims. Known by Greene et al. as the rationalists, these students seek to verify the expertise of authority figures and scrutinise the subjective beliefs held by individuals before accepting them as valid. They also look for corroborating evidence. Although their beliefs in personal justification and justification by authority are likely to be of equal strength, rationalists do not regard the different means of justification as being equally valid in all situations; instead, they adopt what are termed adaptive justification beliefs. Justification beliefs are judged to be adaptive if a person's beliefs and ways of thinking about knowledge and knowing are advantageous. Since different disciplines, and even different topics within the same discipline, can require students to adopt different beliefs and ways of thinking (Sandoval et al., 2016), adaptiveness is domain and potentially topic-specific (Ferguson & Bråten, 2013).

This research is particularly concerned with what makes for adaptive justification beliefs when Science FY students construct arguments using information found in multiple documents. Ferguson et al.'s (2012) Justification for Knowing model has been selected because it operates in the intersection between students' justification beliefs, argumentation, and multiple document comprehension. This model features the two dimensions, personal justification, and justification by authority, found in Greene et al.'s (2008) model and adds a third: justification by multiple sources. Beliefs in justification by multiple sources are thought to be the most adaptive for multiple document comprehension (Ferguson & Bråten, 2013). This is because students with such beliefs have better

sourcing skills, look for corroborating evidence within multiple documents, and have greater engagement in argumentation (Bråten et al., 2014).

The three models of epistemic cognition and associated empirical findings described here indicate that Science FY students are likely to differ in their justification beliefs. It is expected that few, if any, of the researched students will have beliefs in simple and complex knowledge because students entering university are expected to have already passed through what Kuhn (1991) termed the absolutist, and what Greene et al. (2008) termed the realist stage of development. Those who are multiplists may initially accept all claims as being equally valid. Those who are sceptics are likely to seek to justify claims through personal experiences. Such beliefs are maladaptive within scientific communities because scientists rely on the findings of scientific research rather than personal experiences when engaging in argumentation (Duschl & Osborne, 2002). Those who are dogmatists will attend to those with expertise. Having beliefs in justification by authority is thought to be advantageous when students lack prior knowledge (Bråten et al., 2008), which is likely to be true of Science FY students because many are learning a new discipline. Although it is unlikely that any of the researched students will have already become what Kuhn (1991) termed evaluativists and Greene et al. (2008) called rationalists, because these stages of development are typically associated with students at later stages of their undergraduate study, the epistemic cognition models adopted here predict that such students would be able to assess conflicting perspectives with relative ease.

3.2.4 Summary

The models introduced here underpin the research described within this thesis. Toulmin's (1958) model of practical arguments has been adopted as the model of argumentation because its rhetorical framework enables scientists to relate new findings to existing conceptual frameworks (Duschl & Osborne, 2002). The adoption of the two-step validation model (Richter & Maier, 2017) and the CAEM (List & Alexander, 2017) highlights the need to promote high levels of interest and motivation alongside the development of effective cognitive habits when seeking to support students who are reading multiple documents. Two of the models of

epistemic cognition described here, Kuhn's (1991) model of informal reasoning, and Greene et al.'s (2008) Epistemic and Ontological Cognitive Development model, enable the researched students to be described in relation to their justification beliefs. Along with the Justification for Knowing model (Ferguson et al., 2012), these models predict that Science FY students who are dogmatists are likely to find that their justification beliefs are adaptive whereas those who are sceptics or multiplists may find their beliefs to be maladaptive.

3.3 Why do Science Students Need to be Able to Engage in Argumentation?

Argumentation is important for science students because knowledge production within scientific communities is discursive in nature (Stadtler, 2017). This section argues that students studying science at university need to be able to engage in argumentation to develop their understanding of the epistemic basis of science, become part of their scientific communities and contribute to societal debates about controversial socio-scientific issues. It also draws attention to the role that argumentation plays in enabling students to develop deeper knowledge (Sawyer, 2008).

Kuhn (1993) explained that students who are studying science at university need to learn how to construct arguments because scientists are expected to be able to put forward claims, analyse and evaluate evidence, provide warrants, and discuss alternative explanations. Science students who engage in argumentation are therefore better equipped to participate in scientific communities (Rouet & Britt, 2014). Science students become exposed to the argument-like structure of scientific discourse when they engage with the scientific literature (von der Mühlen et al., 2019), where they will encounter claims that are supported by evidence. Consequently, students who understand the underlying structure of arguments will be better able to understand and evaluate scientific information in written documents (von der Mühlen et al., 2009).

Given its potential for exposing students to ways in which knowledge claims are justified in science, Osborne et al. (2004) consider argumentation to be a critically important epistemic task. Yet teaching about the epistemic basis of belief in science is a marginalised aspect of science education for school-age students

(Osborne et al. 2004). These researchers sought to enhance the quality of argumentation by working with science teachers to develop teaching materials and methods for 12-13-year-olds, which they evaluated through conducting a design experiment. In this research, design-based research was used to design and evaluate an approach to supporting university science students.

Beyond the benefits of engaging in argumentation for an individual, the ability to engage effectively in argumentation could also potentially benefit wider society. It is generally agreed that science education in the 21st century should seek to enable active participation in societal debates about socio-scientific issues (Andriessen & Baker, 2014; Guilfoyle et al., 2021, Osborne et al., 2013; Simon et al., 2006). Being able to actively engage with socio-scientific issues requires an understanding of the nature of a scientific argument and its limitations (Simon et al., 2006). Researchers have long been making the case for equipping young people with the skills and proficiencies needed to participate in debates about critical scientific issues (Andriessen & Baker, 2014). Efforts to raise public scientific literacy have typically been directed at secondary schools. As will be discussed shortly, these efforts may have been thwarted by science teachers' pedagogical practices and by subject curricula.

When students evaluate and revise scientific arguments, they are required to relate new data to existing models and theories (Duschl & Osborne, 2002). As such, engaging in argumentation provides opportunities for students to develop deeper disciplinary knowledge (Guilfoyle et al., 2021). Arguments are often constructed using multiple documents. Learning from multiple documents is thought to promote knowledge creation and transformation (Wiley & Voss, 1996; Litman et al., 2017) and lead to deeper understanding (Bråten & Strømsø, 2010; Rouet and Britt, 2014). This is because when students engage with multiple documents, they are required to construct their own understanding through inquiry, interpretation, critique, and evaluation of their sources (Litman et al., 2017).

Studies by Wiley and Voss (1996), Le Bigot and Rouet (2007), and Bråten and Strømsø (2010) have all demonstrated that constructing arguments from multiple documents promotes deeper learning. Wiley and Voss found that undergraduate

students who were asked to write arguments produced more critical analysis than students who were asked to write an account or narrative. Students in the argument-writing condition also used more connectives in their writing, suggesting that they had developed a deeper understanding of the issue. Wiley and Voss concluded that engaging in argumentation moves students from being knowledge tellers to becoming knowledge transformers. Similarly, students in the argument condition of Le Bigot and Rouet's study wrote significantly longer essays than those in the summary condition, used more causal and consequence connectives and included more transformed information, all of which are thought to be indicative of deeper processing. Bråten and Strømsø opted to use an inference task rather than an essay to assess whether writing an argument leads to deeper processing of information than producing a general overview. Nevertheless, their findings indicated that engaging in argumentation had enabled students to construct more integrated mental representations of the documents.

3.4 Why do Some Science FY Students Find it Hard to Construct Arguments from Multiple Documents?

One of the main objectives of this research is to identify the challenges that Science FY students face when trying to construct arguments from multiple documents. Hence the first research question, *RQ1***,** is:

RQ1: What challenges do Science FY students face when constructing arguments from multiple documents?

The reviewed literature indicates four factors should be taken into consideration when seeking to address this research question. Firstly, most students starting university will have had scant opportunities to explore conflicting perspectives during secondary school science classes (Duschl & Osborne, 2002) and will lack experience of using multiple documents, whether printed or online, as resources for producing new knowledge (Litman et al., 2017). Secondly, they are unlikely to have well-developed disciplinary knowledge or hold domain-specific beliefs. This is particularly pertinent for Science FY students because knowing what constitutes

domain-specific beliefs about knowing is especially important when changing disciplines (Green & Hood, 2013). Thirdly, students' justification beliefs affect their ability to engage in argumentation (Bråten et al., 2014) and comprehend multiple documents (Strømsø et al., 2011). The fourth factor is topic interest. The models of multiple document comprehension adopted in this research consider affective engagement and motivation alongside cognition. Yet the importance of topic interest in multiple document comprehension remains contested. For instance, Bråten et al (2018) concluded that time spent on task matters more than students' motivational state. The position taken in this research is that topic interest plays an important role in encouraging students to engage in time-consuming and effortful cognitive processes, and that having high levels of interest in a topic may help students to sustain these efforts when they are working with multiple documents over an extended period.

Prior educational experience

School-age students are rarely, if ever, engaged in the type of dialogue found within authentic research communities. This is because secondary school science classrooms differ markedly from the discursive nature of knowledge production that characterises scientific communities (Litman et al., 2017). Instead, secondary school science teaching often involves the transmission and recall of content (Duschl & Osborne, 2002). Venville and Dawson (2010) proposed that teacher-centred pedagogical practices dominate science classrooms because secondary school science teachers do not typically undertake original research. When this lack of first-hand experience of engaging in argumentation is coupled with pre-service teacher training that aims to ensure future teachers know and understand key scientific concepts, models of science teaching that privilege the science teacher may be reinforced (Duschl & Osborne, 2002).

Another powerful influence on UK secondary school science teachers' practice is the Science National Curriculum. This statutory guidance, which is produced by the Department for Education (DfE) and mandates what must be taught in English classrooms, makes only implicit references to argumentation (Erduran et al., 2022). The word argument or variants thereof does not feature

within the document for 11-14-year-olds (DfE, 2013) and only features once in the guidance for 14-16-year-olds (DfE, 2014). In contrast, disciplines such as Religious Education (RE) have more overt references to argumentation in their syllabi (Erduran et al., 2022).

Even when students engage in argumentation in other disciplines, they are unlikely to be familiar with all components of an argument. For example, Chan et al.'s (2021) analysis of UK RE syllabi found that the claim, rebuttal, and qualifiers were emphasised, but the evidence, warrant, and backing were somewhat neglected. Litman et al. (2017) concluded that students need to be taught what is required for argumentation in science even if they have engaged in argumentation in other disciplines. Students from other countries are likely to face similar gaps in their knowledge. For instance, von der Mühlen et al. (2019) noted that German undergraduate students find the warrant of an argument to be particularly problematic.

While many Science FY students have been educated in UK schools, others have attended schools in their home countries. Research has indicated that the UK is not alone in privileging content acquisition over the consideration of alternative perspectives within secondary school science classrooms. For example, Gil et al. (2010) commented that the Spanish educational system provides limited opportunities for students to gain experience in writing argument essays. Moreover, observations made by Litman et al. (2017) in American schools indicated that science lessons provided children aged 12 to 18-years with very few opportunities to engage in argumentation and they rarely used multiple documents.

Disciplinary knowledge and domain-specific beliefs

It is now widely accepted that students' prior knowledge interacts with argumentation, multiple document comprehension, and epistemic cognition (for example, Mason & Boldrin, 2008; Maier & Richter, 2013; Strømsø et al., 2017). The two-step validation model predicts that students with higher levels of prior knowledge are more likely to engage in the effortful processing of belief-inconsistent information (Richter & Maier, 2017). This position is supported by Trevors et al. (2016), who demonstrated that students with higher levels of prior

knowledge are better able to detect discrepancies. It is also likely that having higher levels of prior knowledge better enables students to engage in argumentation. This is because students who are more familiar with a discipline will better understand how typical and atypical arguments are structured within that discipline (von der Mühlen et al., 2019).

Having domain-specific beliefs is also thought to be important when engaging in argumentation. Domain-specific beliefs are beliefs about what constitutes a knowledge claim and what is considered appropriate evidence within that discipline (Driver et al., 2000). Engaging in argumentation enables students to develop their understanding of what counts as knowledge in a discipline (Guilfoyle et al., 2021). Those who develop their disciplinary knowledge through reading multiple documents are also likely to learn that scientific claims are supported by evidence (von der Mühlen et al., 2019). Once students have developed domain-specific beliefs, they will be able to recognise different perspectives and be able to relate evidence and explanations to claims (Duschl & Osborne, 2002), both of which are important aspects of argumentation in science.

Students' justification beliefs

Justification beliefs influence argumentation and multiple document comprehension (Bråten & Strømsø, 2010; Gil et al., 2010). Beliefs in justification by multiple sources are thought to be the most adaptive and beliefs in personal justification the least adaptive. Beliefs in justification by authority occupy the intermediate position in the hierarchy, though where students lack prior knowledge of complex issues, beliefs in justification by authority are thought to be particularly helpful (Bråten et al., 2008).

Studies conducted by Ivar Bråten, Leila Ferguson, Helge Strømsø, and their collaborators have provided evidence in support of this hierarchy. For example, research conducted in 2011 by Strømsø et al. demonstrated that students with personal justification beliefs perform less well than those with beliefs in authority when studying multiple documents. Having found that students with personal justification beliefs trust documents less and attend to their content less than students with beliefs in justification by authority, Strømsø et al. concluded that

those with personal justification are unable to construct representations of the author's meaning. Ferguson and Bråten (2013) found that even moderate reliance on personal justification often proved to be a barrier to the acquisition of knowledge. This may be because these students lack the sourcing and corroboration behaviours needed to engage in deeper-level processing, which limits their knowledge gains (Bråten et al., 2014). In contrast, students with beliefs in justification by multiple sources were more likely to include citations in their work and link sources to content. These students were also more likely to produce arguments which integrated information from different documents than those with other justification beliefs (Bråten et al., 2014). Indeed, having beliefs in justification by multiple sources was the strongest predictor of sourcing and argumentation skills. This is because students with strong beliefs in justification by multiple sources pay attention to source information and attempt to synthesise different documents. These effective cognitive habits enable such students to perform better in complex reading and writing tasks than those with other justification beliefs.

Given that few undergraduate students arrive at university having had experience of working with multiple documents (Litman et al., 2017), it seems unlikely that Science FY students will have well-developed sourcing and corroboration behaviours. Even if students have effective cognitive habits, the two-step model validation model (Richter & Maier, 2017) predicts that students will differ in whether they allocate the additional cognitive resources needed to evaluate information that contradicts what they had previously believed to be true (Richter & Maier, 2017). Those with beliefs in personal justification are the least likely to engage in the effortful processing of belief-inconsistent information. There is evidence in support of this assertion; Strømsø et al. (2017) found that first year undergraduates with strong beliefs in personal justification typically placed more trust in documents that matched their pre-existing beliefs than those which were inconsistent. Consequently, students with strong beliefs in personal justification are unlikely to move beyond the initial automated validation processes (Richter & Maier, 2017) and will fail to explore the validity of belief-inconsistent information. In contrast, Richter and Maier (2017) expect those with more adaptive justification beliefs to be motivated to construct justified and defensible arguments.

In seeking to understand the challenges that Science FY students face, it is therefore important to know how they justify their knowledge claims. Thus, the second research question, *RQ2*, is:

RQ2: What are the justification beliefs of Science FY students at the researched university?

Topic interest

Well-established models of multiple document comprehension (MDC), such as the Documents Model, pay little regard to whether students are interested in the documents they are reading. In contrast, more recent models of MDC, such as the two-step validation model and the CAEM propose that even when students have effective cognitive habits, their levels of interest, motivation, and affective engagement determine how well they will attend to source information and construct meaningful mental representations of the content.

Our current understanding of the role of topic interest is somewhat limited because researchers rarely, if ever, allow participants to select topics they are interested in (Primor & Katzir, 2018). Even though researchers have expressed concerns that students may lack motivation to process documents given to them in experimental settings (Bråten et al., 2015), studies typically involve students reading documents about topics chosen by the researchers. Efforts have been made to select topics that are likely to matter to 17- to 25-year-olds, with environmental issues such as climate change and nuclear power featuring in numerous studies (for example, Brandmo & Bråten, 2018; Bråten & Strømsø, 2010; Bråten et al., 2008, 2009; Gil et al., 2010; Strømsø et al., 2011). Potential health risks are another popular choice, with several studies featuring documents about the potential health risks posed by mobile phone use (for example, Bråten et al., 2014, 2015; Ferguson et al., 2012; Strømsø et al., 2013). Upper secondary school and university students have also been tasked with reading documents about the potential benefits and risk of sunlight (Ferguson & Bråten, 2013; Ferguson et al., 2013; Stang Lund et al., 2019).

A further limitation of much of the research into multiple document comprehension is that it usually involves students being given a relatively small number of documents to read in a relatively short amount of time, often less than an hour. In this research, students will be expected to read numerous documents over several weeks and are expected to allocate many hours to doing to. As such, it provides a unique opportunity to explore whether having high levels of topic interest sustains students' efforts, particularly when they are facing multiple demands on their time.

Although further research is needed to understand the role played by topic interest when students select their own issues, previous studies have provided some insights into the relationship between topic interest, prior knowledge, and justification beliefs. For example, Bråten et al. (2009) concluded that university students who have high levels of topic interest but low levels of prior knowledge are more likely to rely on those with expertise to justify their knowledge claims. Having strong beliefs in justification by authority can be adaptive when students are first learning a new discipline (Bråten et al., 2008), but over-reliance on external sources of knowledge can limit knowledge gains. Hence, Bråten et al. (2009) advise that students should engage in reading or informed discussion about issues they are interested in before engaging in complex reading and writing tasks.

Although the directionality of the relationships between justification beliefs, topic interest and prior knowledge has not yet been established, Brandmo and Bråten (2018) hypothesized that students' justification beliefs may be predictive of their levels of topic interest. They expected those with beliefs in personal justification to be less interested in complex or controversial topics. In contrast, they expected those with beliefs in justification by multiple sources to become more interested in the topic as they compare documents.

3.5 How can Students be Supported in Constructing Arguments From Multiple Documents?

The research reviewed in the previous section suggests that Science FY students will arrive at university having had few opportunities to understand the rhetorical structure of arguments or to develop effective sourcing and corroboration behaviours when dealing with multiple documents. Those who rely on personal experiences to justify what they believe to be true, rather than having beliefs in justification by authority or multiple sources, are less likely to trust documents which provide conflicting information and will attend less to their content, leaving them poorly equipped to use information from multiple documents when constructing arguments. If poorly developed cognitive habits and maladaptive epistemic cognition are coupled with low affective engagement, constructing arguments from multiple documents will be immensely challenging. Given that Science FY students are diverse, they are likely to differ in their previous educational experiences, prior knowledge, and justification beliefs. Consequently, it is likely that these students will face different challenges when trying to construct an argument from multiple documents.

One of the main objectives of this research was to design and evaluate an intervention to support these students in constructing arguments from multiple documents. The design aspect of this objective is expressed in Research Question 3 (*RQ3*)**:**

RQ3: What the key design principles of an intervention designed to support students at the researched university in constructing an argument from multiple documents?

The effectiveness of this intervention in achieving its objective will be explored through Research Question 4 **(***RQ4***):**

RQ4: What are students' experiences of and outcomes from this intervention?

3.5.1 The Design Principles

The intervention was informed by five design principles (DPs), which were identified within the reviewed literature. This section introduces these principles and discusses the literature that underpins them.

DP1: Provide students with opportunities to develop their disciplinary knowledge and domain-specific beliefs.

Students will be encouraged to read sources that develop their disciplinary knowledge and enable them to gain awareness of how evidence in support of claims is generated through scientific research. This because having knowledge of the practices within a discipline is thought to support the development of domain-specific beliefs (Kienhues et al., 2016). Students who have ways of thinking about knowledge and knowing that are aligned with those of the discipline they are studying are said to be demonstrating appropriate adaptiveness to context (Sandoval et al., 2016). In this research, having strong beliefs in justification by authority is thought to be adaptive because Science FY students are expected to lack prior knowledge. Having strong beliefs in personal justification is thought to be maladaptive because scientists are unlikely to rely on personal experiences to justify what they believe to be true. Having similar strength beliefs in both personal justification and authority is also likely to be maladaptive because students with such beliefs are not expected to have the sourcing and corroboration skills needed to evaluate alternative perspectives and determine which is the most credible. Without these skills, students are likely to place equal trust in all perspectives; this is maladaptive because scientists are rarely pluralistic.

Although argumentation necessarily exposes students to differing perspectives, it is worth noting that directing students to read texts that differ markedly from their own perspective when first learning about a new discipline is thought to be best avoided. This is because students' pre-existing beliefs influence the extent to which they engage with perspectives that differ from their own (Bråten, 2016). Having even moderate reliance on personal justification is likely to be a barrier (Ferguson & Bråten, 2013). Hence when learning a new discipline,

students may benefit from reading sources that align with their pre-existing beliefs before being guided to engage with those that offer alternative perspectives.

DP2: Allow students to choose topics which they are interested in.

Students will be given the opportunity to construct arguments about a controversial socio-scientific issue that interests them. This differs markedly from what typically happens in MDC research, where the inquiry questions are usually set by the teacher or researcher (Primor & Katzir, 2018). In placing equal importance on affect and cognition, this design principle draws upon the CAEM (List & Alexander, 2017), which predicts that it is only when students are affectively engaged and are enacting effective cognitive habits that they will be able to comprehend multiple documents.

DP3: Teach students how to construct arguments.

Students will be taught about the components of an argument and will learn about the relationships between these components through assessing arguments written by others and engaging in dialogue about these arguments. This design principle is founded on research by Duschl and Osborne (2002), Nussbaum et al. (2008), Osborne et al. (2016) and von der Mühlen et al. (2019). Collectively, these researchers indicated that students need to be taught about the components of an argument, taught how to construct arguments from these components, such as relating evidence to claims, and should discuss what they think about arguments written by others. Care will be taken when teaching students about arguments to avoid giving the impression that students should be putting forward their own views. This is because Bråten et al. (2014) found that permitting students to express their own opinions sometimes impeded their ability to synthesise information from multiple documents.

DP4: Equip students with the cognitive habits needed to engage with multiple documents.

Students will be taught how to use resources that encourage them to consider multiple factors when assessing source credibility and evaluating content information. This design principle draws from the research conducted by Wiley et al. (2009), who developed the SEEK model as a means of encouraging students to consider the type of evidence presented, how well this evidence fits existing explanations of the phenomena, and how well the new information fits with prior knowledge. Students in the SEEK group were better able to distinguish between reliable and unreliable sources of information, were more likely to provide causal accounts, were less likely to mention erroneous causes, and made more references to sources. While Wiley et al.'s findings indicate that providing students with templates can enable them to evaluate and use sources more effectively, care will need to be taken to avoid fostering over-reliance on such devices in students who lack affective engagement because the CAEM predicts such students often adopt simple heuristics (List & Alexander, 2017).

Notwithstanding the valid concerns about creating an over-reliance on templates, the researched students will also be provided with resources that will enable them to record key pieces of information when reading journal articles. This is because they are unlikely to have encountered such sources previously and so it will be also necessary to teach them how to read reports of scientific research effectively and efficiently. These resources will encourage students to consider whether the findings relate to previous knowledge. This should prove beneficial because asking students to pay attention to information that conflicts with their prior knowledge will encourage them to move beyond automated processing of belief-consistent information (Richter & Maier, 2017) and begin engaging in corroboration behaviours.

DP5: Provide positive, timely feedback, which motivates students to engage in effortful processing of belief-inconsistent information.

Efforts will be made to promote a favourable motivational state by providing students with positive feedback. This is because students will only allocate the

cognitive resources needed for elaborative processing of belief-inconsistent information if they are motivated to do so (Maier & Richter, 2014). Students with beliefs in personal justification may especially benefit from receiving positive feedback about the efforts they are making to process discrepant information (Maier & Richter, 2014). Provision of timely feedback is also thought to be important when teaching students about the components of arguments (Simon et al., 2006).

3.6 Summary

Science students need to be able to engage in argumentation if they are to become active participants within their scientific communities. Engaging in argumentation also brings individual and societal benefits. Yet students often struggle to construct arguments from multiple documents. This chapter has reviewed three areas of literature, argumentation, multiple document comprehension, and epistemic cognition, and theoretical models which operate within the intersections between them. Four factors that affect students' ability to construct arguments from multiple documents have been surfaced. These are: prior educational experience, disciplinary knowledge and domain-specific beliefs, students' justification beliefs, and topic interest. The final section of this chapter has introduced the five design principles that will be implemented in an intervention that aims to support students with differing justification beliefs in constructing arguments from multiple documents.

4 Methodology and Methods

4.1 Introduction

This chapter of the thesis is concerned with how research was conducted to understand how to better support undergraduate science students with differing justification beliefs in constructing an argument from multiple documents. This research aim was pragmatic; thus, this chapter begins by describing pragmatism and explaining how this approach aligns with design-based research, the methodology used within this study. Pragmatism is pluralistic, which means it allows for the use of mixed methods; the explanatory sequential research design used to generate data from two successive cohorts of Foundation Science students at the researched university is described. This is followed by a description of the materials used to collect quantitative and qualitative data. As design-based research is an iterative process, the stages of this process and the research questions addressed during each stage are given in the procedure section of this chapter, along with details of how the materials were used within Studies 1 and 2 to collect data. However, specific information such as the nature of the sample and timeline are provided in the relevant study chapters. Next, there are details of the data analyses, including the statistical methods used and the analytical process known as Framework (Ritchie and Spencer, 1994). This chapter concludes with consideration of ethical issues that are pertinent when the researcher is also the practitioner.

4.2 Research Aim

This research has its origins in informal observations made while teaching and supporting students at the researched university. I observed that Foundation Year students were struggling to synthesise information from multiple documents when trying to construct scientific arguments about scientific issues. This is not a satisfactory situation, particularly since the Science FY course aims to prepare students for undergraduate science courses where they will be expected to arrive at critical judgments having summarised and synthesised information from a variety of sources (Quality Assurance Agency [QAA], 2019). Having identified a problem, I

began exploring the literature about argumentation and multiple document comprehension. I soon noticed that these areas of literature often overlap with research about justification beliefs. This study operates within the intersection of these three research areas. Hence, the primary aim of this research is to understand how students with differing justification beliefs can be better supported in constructing arguments using multiple documents. This aim will be addressed through four research questions:

RQ1: What challenges do Science FY students face when constructing arguments from multiple documents?

RQ2: What are the justification beliefs of these students?

RQ3: What are the key design principles of an intervention designed to support these students in constructing an argument from multiple documents?

RQ4: What are students' experiences of and outcomes from this intervention?

4.3 Pragmatism

Pragmatism originated as a philosophical framework in 19th century America (Maxcy, 2003). Pragmatists accept that individuals differ in how they interpret their experiences; these differences arise because individuals differ in their understanding of the nature of the world (Morgan, 2014). This research is concerned with students who differ in how they justify what they believe to be true. It is also concerned with the challenges these students face when trying to construct arguments from multiple documents, and their experiences of an intervention designed to support them in doing so. Pragmatic researchers are often guided by their own beliefs and (Badley, 2003). Here, the design of the intervention was grounded in the literature, but its implementation was guided by my pedagogical philosophy (see Chapter 1).

Popularised during the 20th century as being primarily concerned with practical applications (Crotty, 1998), pragmatism is now considered to be a pluralistic approach to enquiry. This means that it does not privilege one paradigm over another, is not restricted by the expectations of established research

communities and does not adhere rigidly to the traditional triad of ontology, epistemology, and methodology (Badley, 2003; Morgan, 2014). Instead, pragmatists select whichever research methods will allow them to better understand how their participants perceive real-world actions, situations, and consequences (Creswell, 2009; Morgan, 2014). Hence pragmatists view quantitative and qualitative methods as different but equally valid tools for research (Badley, 2003), and are comfortable in adopting mixed methods approaches. By adopting pragmatism, this research was able to use analysis of quantitative data to measure changes at the cohort level and qualitative data to explore the subjective experiences of individuals with differing justification beliefs.

4.4 Design Based Research

Design-based research (DBR) is underpinned by pragmatism. This is because it aims to develop researchers' understanding how learning is supported within real-world learning environments (Cobb et al., 2003). As such, DBR is described as doing 'real work' (Cobb et al., 2003); hence, it is aligned with pragmatist thinkers such as John Dewey (1859-1952) and Kurt Lewin (1890-1947) who were both concerned with how practice and theory inform each other.

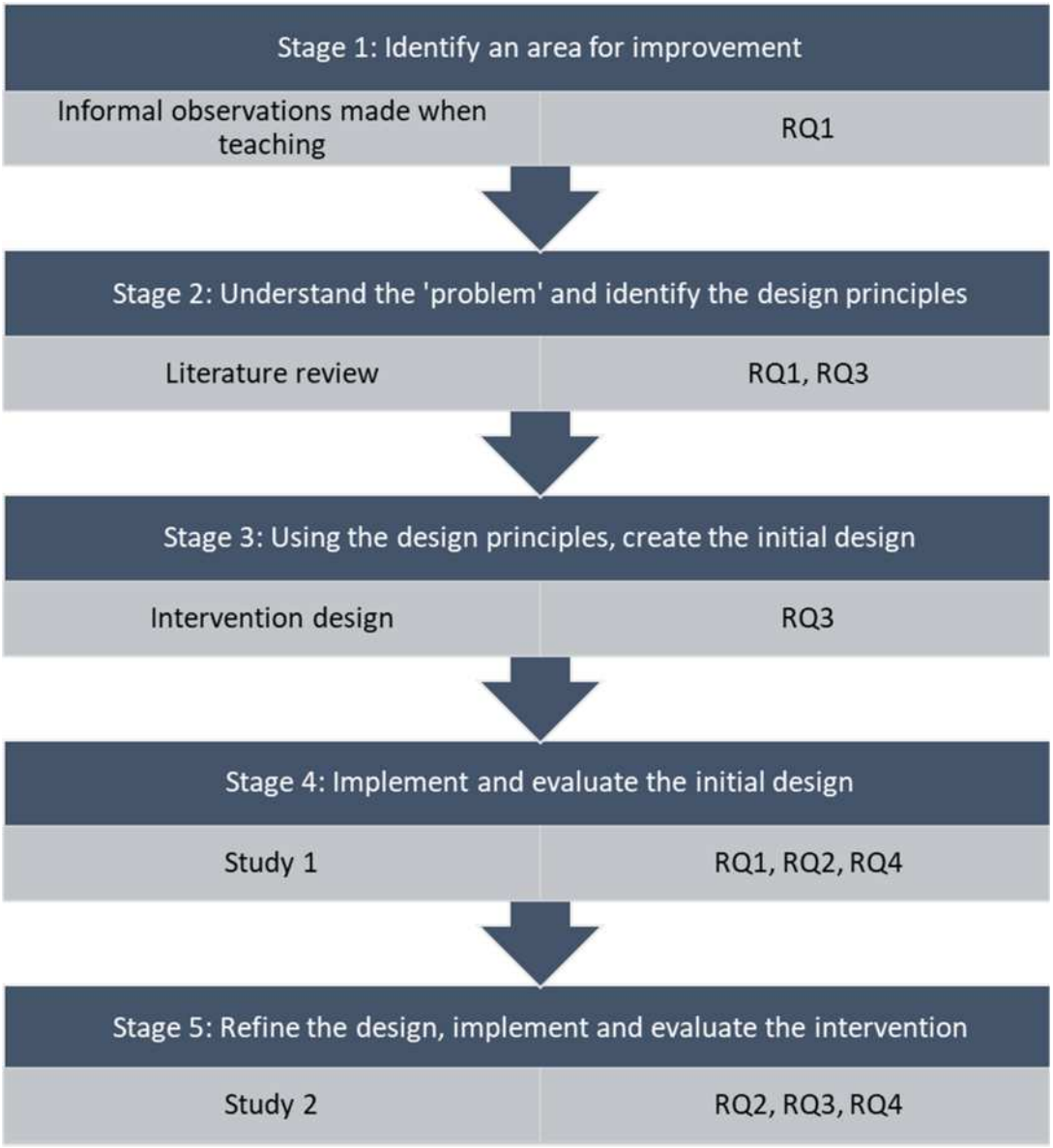
Brown (1992) and Collins (1992) are usually credited with establishing DBR; both were frustrated with other methods of educational research. Brown was specifically seeking to overcome the limited explanatory and predictive power of laboratory studies. Collins identified the need for a methodology which allowed for progressive testing and refinement of design variants. The methodological framework that Brown and Collins devised is iterative and highly interventionist. It begins by identifying a problem, and then proceeds through successive stages of design, implementation, initial evaluation, refinement, further implementation, and evaluation until an acceptable solution is reached. provides a schematic representation of how design-based research was enacted over the course of two studies and indicates the research questions (*RQs*) addressed during each stage of the process. Having identified a problem through informal observations (stage 1), the purpose of the literature review (stage 2) was to identify challenges faced by students when trying to construct scientific arguments from multiple documents

(*RQ1*) and to elicit the design principles (*RQ3*) that would guide the design of the first iteration. DBR is typically associated with humble theories; these are theories that explain and predict how students respond to particular aspects of their learning (Bakker & van Eerde, 2015; Cobb et al., 2003; Collins et al., 2004; Edelson, 2002) and have local impact (Barab, 2006). Here, the models of epistemic cognition, multiple document use, and argumentation identified within the literature review (Chapter 3) guided the development of five design principles. The implementation of these principles during stage 3 of the DBR process was informed by my pedagogical philosophy, which is founded in scaffolding (See Chapter 1). As such, the teaching, learning, assessment, and feedback activities often aimed to support students through the process of constructing an argument from multiple documents by providing them with learning aids, such as checklists and templates, and engaging them in dialogue. The designed activities were embedded into three workshops, an individual tutorial, and coursework assignments within Semester 2 of the Studying Science at University module.

The first iteration of the designed intervention was implemented and evaluated during Study 1. This stage of the process also provided an opportunity to understand students' experiences. Consequently, Study 1 was concerned with better understanding the challenges that students face (*RQ1*), sought to identify the justification beliefs held by this cohort of Science FY students (*RQ2*) and began to evaluate students' experiences of and outcomes from the intervention activities (*RQ4*). As is typical of DBR (Cobb et al., 2003; Collins et al., 2004) analysis of students' experiences of the initial design during stage 4 led to some of the design principles being revised (*RQ3*). The updated teaching, learning, feedback, and assessment activities were enacted in stage 5, when Study 2 took place. The justification beliefs of the second cohort of Science FY students were measured (*RQ2*) and their experiences of and outcomes from the intervention were explored (*RQ4*).

Figure 4-1

The Five-Stage Process of Design-Based Research



4.5 Research Design

The research design needed to allow for iterative phases of data generation and analysis using both quantitative and qualitative methods. The first phase had a one-group pre-test post-test design (Cranmer, 2017). This design is often used by social scientists when evaluating the effectiveness of educational interventions. It has a linear order, with a pre-test assessment of the dependent variable occurring first, followed by the treatment, and then by the post-test assessment. All participants complete the same pre- and post-test assessments and are subjected to the same treatments. The effectiveness of the treatment is measured by calculating the

difference between scores attained in the pre-test and post-test assessments. Here, all participants completed a questionnaire about their justification beliefs before the intervention took place. They then experienced the teaching, learning, and feedback activities of the designed intervention, which were embedded into Semester 2 of the Studying Science module. All participants completed the post-intervention questionnaire after submitting their Semester 2 assignments. The differences in the pre- and post-intervention scores were calculated and used as a measure of the changes to justification beliefs.

The pre-intervention scores for justification beliefs, and the changes to these, were also used to identify a small subset of participants. Details of the selection process for these participants follow shortly. These participants were invited to retrospective interviews. Qualitative analysis of the interviewees' reported experiences aided the interpretation of the quantitative data. Hence this aspect of the study had a sequential explanatory design (Creswell, 2009). The explanatory purpose differed between the two studies. In Study 1, participants' reported experiences enabled me to understand whether students with justification beliefs face different challenges when trying to construct arguments from multiple documents (*RQ1*). The Study 1 participants were also asked about their experiences of the intervention, but their responses were primarily used as a means of refining the design principles and implementation thereof *(RQ3*). In Study 2, the qualitative data generated by the interviews enabled me to explore the effectiveness of the intervention activities more fully and compare the experiences and outcomes of those with different justification beliefs (*RQ4*).

4.6 Sampling and Participants

Studies 1 and 2 were conducted at the university where I am employed as the Course Director for the Science FY programme. Students from this programme, who will be referred to as Science FY students, were asked to consent to being participants in the study. The Study 1 and Study 2 cohorts are described in further detail within Chapters 6 and 7 respectively.

In both Study 1 and Study 2, participants could only be considered for the retrospective interview if they had consented to do so. It was also necessary for the

audio recording of their individual tutorial to be available, along with their reflective commentary. Almost all who consented to be interviewed were available for selection. The selection process was systematic, with participants firstly being sorted based on the scores for their pre-intervention justification beliefs using a statistical method known as cluster analysis. Two or three individuals from each cluster, who had different magnitudes and directions of change to their justification beliefs from pre- to post-intervention, were invited to participate in retrospective interviews. This systematic approach to sampling ensured that the interviewees were heterogenous, which was important given that the key aim of this research was to understand how best to support students with differing justification beliefs.

4.7 Materials

4.7.1 Questionnaire

In common with other multi-dimensional models of epistemic cognition, the Justification for Knowing model (Ferguson et al., 2012) is associated with a self-report measure. The Justification for Knowing Questionnaire ([JFK-Q], Ferguson et al., 2013) is a 14-item questionnaire which measures justification beliefs about knowing on three sub-scales: beliefs in personal justification; beliefs in authority and beliefs in multiple sources. This questionnaire can be found in Appendix A. There are three items relating to beliefs in personal justification, all of which relate to one's personal views or opinions; for example, '*What is a fact in the natural sciences depends on one's personal views*'. The beliefs in authority sub-scale includes six items, two of which refer to scientific research, for example, '*I believe in claims that are based on scientific research*'; two to science teachers, for example, '*If a science teacher says something is correct, then I believe it*'; one to scientists, '*If a scientist says that something is a fact, then I believe it*' and one to published materials, '*Things that are written in natural sciences textbooks are correct*'. Beliefs in multiple sources are measured by five items, all of which refer to checking several sources; for example, '*I can never be sure about a claim in the natural sciences until I have checked it with at least one other source*'. Participants recorded their responses to these items using a ten-point Likert scale, where 1 was disagree completely and 10 was agree completely.

Ferguson et al. (2013) reported Cronbach's alpha values ranging from 0.57 to 0.81 when they assessed the internal consistency of the three sub-scales within the JFK-Q. In this research, the Cronbach's alpha values ranged from 0.48 to 0.88 in Study 1, and from 0.72 to 0.80 in Study 2. Table 4-1 shows the values for the three sub-scales for the pre- and post-intervention JFK-Q in Studies 1 and 2.

Table 4-1 *Cronbach's Alpha Values for the Three Sub-Scales of the JFK-Q in Studies 1 and 2 at Pre- and Post-Intervention*

Sub-scales	Study 1 (*N*=25)		Study 2 (*N*=30)	
	Pre-Intervention	Post-Intervention	Pre-Intervention	Post-Intervention
Personal Justification	0.56	0.67	0.73	0.78
Authority	0.88	0.70	0.79	0.80
Multiple Sources	0.48	0.71	0.72	0.76

The Cronbach's alpha values for the personal justification and multiple sources sub-scales in the pre-intervention questionnaire in Study 1 fell below the threshold value of 0.65, which is generally agreed to signify a satisfactory level of internal consistency. The value of 0.56 for the personal justification sub-scale was, however, comparable to the reliability estimates obtained by the questionnaire's designers, Ferguson et al. (2013). Since the subsequent calculations of Cronbach's alpha for all sub-scales exceeded the threshold value, this questionnaire was retained as the primary means of measuring participant's justification beliefs.

4.7.2 Interviews

The decision to conduct semi-structured retrospective interviews was influenced by Sandoval et al. (2016), who noted that this form of data collection allows individuals to surface those aspects of epistemic cognition that they are not otherwise consciously aware of. Although the aim of the interviews was to enable participants to describe their experiences, the extent to which participants were asked to reflect upon the effectiveness of the intervention differed between Study 1 and Study 2.

The Study 1 interview began with participants being invited to talk about the controversial socio-scientific issue they had chosen as the basis of their argument and explain why they had chosen it. If required, they were prompted to describe what they already knew about the issue, whether they had strong views about the issue and whether these views had changed during the assignment. They were also asked if they were interested in what experts thought about this issue. Their responses to these questions were used when creating their vignettes. The interviewed participants were then directed to think back to the teaching, learning, and feedback activities that had taken place before they had started the Making an Argument assignment. They were invited to discuss which activities they had found most helpful and which were least helpful in supporting them in constructing their argument from multiple sources. The final part of the interview included questions about their sourcing and corroboration behaviours, with participants being asked specific questions about their experiences of checking an author's credentials and engaging with contradictory sources. They were also asked if there was any advice they would give to future Science FY students about how to approach this piece of work. These questions enabled me to identify challenges that the interviewed participants had faced (*RQ1*). Their responses to these questions brought about revisions to the design principles and their implementation ahead of the second iteration of the intervention (*RQ3*). The protocol for the Study 1 interviews can be found in Appendix B.

In Study 2, the interviews shifted from being focused on participants' individual experiences of constructing an argument from multiple sources and instead placed much greater emphasis on their experiences of the teaching, learning, and assessment activities that had been designed to support them through the process. The interview began with students being asked whether the activities within the workshop about a controversial socio-scientific issue had helped them to understand that there can be different perspectives about an issue. They were also asked to evaluate whether the workshop activities had helped them choose their own issue. The probes for this question explored participants' prior knowledge and pre-existing views of the issue they had chosen, with their responses to these questions being used when creating the vignettes. During the second part of the

interview, students were reminded of the workshops about how to construct and refine arguments. They were then asked to describe what they had learnt, discuss whether evaluating arguments written by past students had proved helpful, and invited to make suggestions about how these workshops could be improved. The final part of the interview asked students to consider what they had learnt from the workshop that aimed to teach students how to critically evaluate source and content information. All aspects of this interview enabled me to explore the effectiveness of the intervention (*RQ4*). The protocol for the Study 2 interview can be found in Appendix C.

4.7.3 Individual Tutorials

The designed intervention featured an individual tutorial of 15-20 minutes duration. In Study 1, these were designed to provide informal feedback about the progress that students were making in completing the Making an Argument assignment. Students were advised that by the time of their tutorial they should have selected a controversial socio-scientific issue and begun to read and evaluate sources relating to this issue. The tutorial also offered the opportunity for students to ask questions about the assignment or share draft work for informal feedback. In Study 2, students were required to submit annotated sources ahead of the tutorial. The first part of the tutorial was given over to providing feedback on these sources and the annotations, with the remainder of the time being used to answer questions about the assignment. The tutorials were audio recorded to allow students to listen back when writing the reflective commentary assignment, which is described below.

4.7.4 Reflective Commentaries

The Reflective Commentary assignment was a compulsory coursework assignment in which students used the Gibbs (1988) structured debriefing framework to reflect on the feedback they were given in the individual tutorial and the study strategies they employed while completing the Making an Argument assignment. The Gibbs framework firstly asks students to describe their feedback and articulate what they thought and felt about this; this is termed their response. They are then asked to evaluate, by identifying the positive and negative aspects of their feedback. The

next stages are to analyse and conclude; this involves making sense of the feedback by considering their own actions. The final stage is action planning; here students consider what actions to take in response to the feedback they were given. The Reflective Commentary assignment guides students through these six stages by asking them to answer three questions:

Question 1

Describe the feedback that you were given during the individual tutorial. In your response, you should comment on what you thought and felt about this feedback. You should also discuss the feedback you received about things you had done well and the areas for improvement.

Question 2

Thinking about the feedback you received, how well do you think you carried out each of the following processes:

- Synthesising information from sources?
- Providing evidence in support of your claims?
- Critically evaluating evidence?

Make sure that you comment on your own actions

Question 3

In the light of the feedback you received during the individual tutorial, what is your action plan for completing the 'Making an argument' assignment? Use examples wherever possible when describing and explaining your action plan. It is expected that your actions will relate to the issues you discussed in your answers to questions 1 & 2.

4.8 Procedure

The one-group pre-test post-test design was enacted through a linear ordering of the measurement and treatment activities. The first measurement activity, the pre-intervention JFK-Q was completed in late January; this is the start of the second

term of the researched university's academic year. Students were introduced to the research study, given information about consent, and asked if they would be willing to participate in the study. All students were given consent forms to sign, but only those who signed the forms were given a printed copy of the JFK-Q to complete. The participants were asked to complete the questionnaire individually, read each item carefully, and remain silent until everyone had finished.

The treatment activities were embedded into Semester 2 of the Studying Science module. In Study 1, the designed intervention featured three workshops; Study 2 had four workshops. These were scheduled in consecutive weeks in February. The individual tutorials followed later in the term, taking place during a three-week period starting in early March.

The second measurement activities took place after participants had completed the Making an Argument and Reflective Commentary assignments. The post-intervention questionnaire was completed as an in-class pen and paper activity in late April in Study 1 and in late March in Study 2; the timing varied due to the Easter break. The retrospective interviews took place after students had completed their summer examinations. In Study 1, the first interview took place in late May and the final interview took place in mid-June. In Study 2, all interviews were conducted during a week in mid-June. All interviews took place in small meeting rooms and lasted around 30 minutes. They were audio recorded using a monaural boundary microphone connected to a tablet computer. The audio recordings were captured using the tablet's voice recorder application and transcribed into Word files that, along with the audio recordings, were stored on password-protected storage drives.

4.9 Analysis

4.9.1 JFK-Q Scores

The scores for each sub-scale of the JFK-Q were calculated by adding together the ratings given for each item of the sub-scale and then dividing by the number of items. These values provided a composite measure of each participant's beliefs in personal justification, authority and multiple sources and enabled four key analyses. Firstly, the scores from the pre-intervention JFK-Q were used to sort participants based on their pre-intervention justification beliefs. The sorting process was conducted using a statistical technique known as cluster analysis, which is described below. Secondly, the differences between the pre- and post-intervention JFK-Q scores were calculated. Knowing these differences enabled participants from within the same cluster who had different magnitudes and directions of change to be identified as potential interviewees. The differences between pre- and post-intervention JFK-Q scores also provided an important outcome measure. The final key analysis involved the creation of post-intervention clusters. The movements of participants between their pre- and post-intervention clusters were tracked to identify different trajectories of change.

Cluster Analysis

Given the heterogenous nature of Science FY students, the participants were expected to have different justification beliefs. Justification beliefs are thought to affect how individuals manage multiple documents (Bråten & Strømsø, 2010) and the extent to which they engage in argumentation (Gil et al., 2010). It was therefore important to identify individuals with different justification beliefs when selecting potential interviewees. Cluster analysis provides an objective means of sorting heterogenous samples into groups (Hair et al., 2014). Although not homogenous, individuals within the same cluster have features in common. The common features of individuals within the same cluster differ from the common features of individuals within a different cluster. For example, in Ferguson and Bråten's (2013) study, individuals in one cluster had moderate scores in all three justification beliefs whereas those in another cluster had low personal justification

beliefs and moderate beliefs in authority and multiple sources. Hence the key difference between these clusters was their beliefs in personal justification.

Cluster analysis can be conducted using different methods. This research uses Ward's method, which utilises the Euclidean distance between each pair of participants as a means of determining which participants share similar characteristics. The Euclidean distance is calculated by taking the square root of the sum of the squared differences between two participants. Ward's method seeks to minimise the increase in the total sum of squares when clusters are combined (Everitt, 2011). Ward's method is sometimes criticised for being more sensitive to outliers than other hierarchical agglomeration methods. This was, however, of benefit within this research because its aim was to understand how to support students with differing justification beliefs. It was therefore important to identify individuals whose beliefs differ from the 'norm' such that their experiences could be better understood.

Hierarchical agglomerative methods initially place individuals in their own cluster before combining them with other individuals to form clusters of increasing size. All individuals ultimately end up in the same cluster. Non-statistical methods were used to select the appropriate numbers of clusters. The first of these methods involved visual inspection of dendrograms, which are graphical representations of the distances between clusters. The dendrograms from Study 1 can be found in Appendix H and those from Study 2 in Appendix I. If splitting a large cluster into two or more smaller clusters enabled 'outliers' to be identified, the solution with more clusters was adopted. This is because it was important to identify individuals whose justification beliefs differ from others so that they could be invited to participate.

When naming the clusters, the relative strengths of the scores for beliefs in personal justification and authority were compared to those described within Greene et al.'s (2008) model of cognitive development. It was sometimes also necessary to consider the alignment between the justification beliefs expressed by the interviewees and those described within Kuhn's (1991) developmental model of informal reasoning.

Testing for Differences in Pre- and Post-Intervention Scores

The pre- and post-intervention JFK-Q scores were subjected to statistical analysis to establish whether these scores had changed significantly over the module. This was one of the outcome measures (*RQ4*). The Wilcoxon matched-pairs signed-ranks test was chosen as being the most suitable statistical test. This is because the JFK-Q scores were obtained using a Likert scale, which asks participants to rank their responses, meaning that the data are ordinal. Ordinal data cannot be analysed using statistical tests that assume the Normal distribution. The Wilcoxon matched-pairs signed-ranks test is a non-parametric test that compares values obtained from the same participant before and after an intervention has taken place. It does so by comparing the signs of ranked values. Where there was no difference between the value of the pre- and post-intervention scores, Pratt's signed ranks zero procedure method was used. It is thought that this approach is better than removing tied ranks from the sample when analysing Likert-scale data because it does not alter the Type I error rate (Derrick & White, 2017). That is, it does not alter the likelihood of claiming there is a statistically significant difference when such a difference does not exist.

It was hypothesised that if the designed activities were effective in promoting the adoption of adaptive justification beliefs, participants would report weaker beliefs in personal justification and stronger beliefs in justification by multiple sources in the post-intervention JFK-Q. Beliefs in justification by authority were also expected to strengthen. As the direction of change was predicted for each justification belief, the statistical analysis was conducted using one-tailed tests. One-tailed tests are more powerful, which means they are more likely to detect a statistically significant difference if one exists.

In addition to comparing the pre- and post-intervention scores for each of the justification beliefs, the change to a summed score of justification beliefs was subjected to statistical analysis. To create a summed score, the values for beliefs in personal justification needed to be transformed. This was because a low value for these beliefs was deemed adaptive. A simple calculation, max - x, where x was the original score reported by the participant and max was the maximum possible score, was used to transform the values for beliefs in personal justification. The

resulting value was added to the scores for beliefs in authority and multiple sources to create a summed score. The summed scores before the intervention were then compared to those after the intervention using the Wilcoxon matched-pairs signed-ranks test. As all the scores were now on scales where high values were more adaptive, it was hypothesised that the summed score would increase if the intervention was effective.

Since the summed scores used data that had previously been subjected to statistical analysis when comparing the changes from pre- to post-intervention for each justification belief, it was necessary to adjust the significance level (α) to account for multiple comparisons when assessing the statistical significance of the change to summed scores. The Bonferroni method was used. This involves dividing α by the number of tests, *m*, conducted using the same data. Here, α was 0.05 and *m* was 2, hence the adjusted significance level was 0.025. Hence the assessment of changes to the summed scores was more conservative than that for the separate justification beliefs.

4.9.2 The Framework Analytical Process

The findings for Research Question 1 and Research Question 4 were analysed using an analytical process known as Framework, which was originally developed by Ritchie and Spencer (1994) as a means of managing and analysing qualitative data gathered during applied policy research. Framework involves five distinct though interconnected stages: familiarisation; identifying a thematic framework; indexing; charting; and mapping and interpretation (Ritchie & Spencer, 1994).

During the familiarisation stage, I became immersed in the data and began to note recurrent ideas by listening to all the audio recordings from the individual tutorials and retrospective interviews and reading the accompanying transcripts. I also read all the reflective commentaries written by the interviewed participants. Next, I used the notes created during the familiarisation stage to generate a thematic framework and construct matrices into which data were entered and indexed. These initial matrices were dominated by *a priori* themes, which is a typical feature of Framework (Ritchie and Spencer, 2011).

Table 4-2 *Initial Thematic Framework for Research Question 1 in Study 1*

Theme	Subtheme
1. Comprehending multiple documents	1.1. Sources 1.2. Reading 1.3. Corroboration 1.4 Contradictory information 1.5 Uncertainties 1.6 Critical evaluation
2. Engaging in argumentation	2.1 Contested issue 2.2 Argument structure 2.3 Arguing
3. Using multiple documents to construct arguments	3.1 Making and supporting claims 3.2 Using data 3.3. Using critically evaluated sources 3.4 Reaching a conclusion
4. Prior experience, knowledge, and beliefs	4.1 Prior experience 4.2 Knowledge of specific terms 4.3 Statistical analysis 4.4 Knowing conventions 4.5 Feeling part of the community
5. Dealing with negative affect	5.1 Underprepared 5.2 Overwhelmed 5.3 Being cautious 5.4 Being unsure / seeking assurances 5.5 Maintaining intrinsic motivation
6. Time-management	6.1 Managing own time 6.2 Procrastinating 6.3 Prioritising 6.4 Having time off
7. The writing process	7.1 Grammatical structures 7.2 Turning thoughts into text 7.3 Sticking to the word limit

For example, in Study 1 the initial thematic framework generated for *RQ1*, which is shown in Table 4-2 was dominated by themes that had been identified in the literature, such as the difficulties that students face when trying to comprehend multiple documents.

Having created an initial thematic framework, data from the transcripts and reflective commentaries were indexed then against themes and sub-themes. Ritchie and Spencer (2011) advise that category labels should be relatively short and

remain largely descriptive to enable data to be retrieved easily during the later stages of the process. For example, when indexing data during Study 1 for Research Question 1, sub-theme category labels were either single words such as 'reading' and 'corroboration', or short phrases, such as 'making and supporting claims' and 'being cautious'. These category labels, alongside numerical values that represented the hierarchy between themes and subthemes, were then used to annotate sections of text within transcripts and the reflective commentaries.

The fourth stage, which is known as charting, involves organising the indexed data into a table. In this research, data were organised by cluster and by changes to beliefs in personal justification. Associations made when charting data often led to changes being made to the thematic framework, with themes and sub-themes being re-organised, combined or even removed. For example, when mapping and interpreting the data for Research Question 1 in Study 1, it became apparent that procrastination, which had been categorised within the time-management theme, was a maladaptive response to experiencing negative affect. It also became apparent that other sub-themes within the time-management theme did not occur independently of the other challenges described by participants. For example, many of the instances of 'Managing own time' co-occurred with the challenges of managing multiple documents. Consequently, time-management was no longer conceptualised as being an independent theme but was instead viewed as either being a contributing factor or response to other difficulties that participants faced. Other revisions during this stage included creating a new 'umbrella' theme of 'Cognitive Challenges', with Themes 1, 2, and 3 as described in Table 4.2 becoming its three sub-themes. The final list of themes for Research Question 1 in Study 1 is shown in Table 4-3.

Table 4-3 *Final Thematic Framework for Research Question 1 in Study 1*

Theme	Subthemes
1. Having limited domain-specific understanding	
2. Cognitive challenges	2.1 Managing multiple documents
	2.2 Engaging in argumentation
	2.3 Integrating information from multiple documents to construct arguments
3. Having maladaptive strategies for dealing with negative affect	

The final stage of Framework is known as mapping and interpretation. This involves comparing participant responses, looking for patterns, and seeking explanations (Swallow et al., 2003). As one of the key aims of this research was to understand whether participants with different justification beliefs face different challenges, this stage of the process involved comparing the experiences of those with the same and different justification beliefs. Chapter 6 describes how the themes identified in relation to Research Question 1 in Study 1 were mapped across the interviewed participants.

4.10 Ethics

This study is an example of practitioner research, and as such ethical concerns were paramount as I juggled the dual roles of practitioner and researcher. This section will therefore firstly explain the importance of thinking ethically when both a practitioner and a researcher. It will then deal with the procedural aspects of ethics by describing how ethical approval was sought.

4.10.1 Ethicality and Axiology

The reciprocal relationship between research and practice is both fruitful and problematic; having context-bound knowledge can aid understanding, but

difficulties arise when the boundaries between practitioner and researcher become blurred (Cochran-Smith & Lytle, 2007). I therefore made deliberate efforts to present myself as a doctoral candidate when asking students to participate. Notwithstanding this, it is probable that some students chose to not opt out from participating rather than actively deciding to participate (Norton, 2007). Moreover, it was sometimes tricky to avoid exceeding the rights of access to participants (Fraser, 1997). For example, when participants were absent from sessions where data collection was taking place, it was difficult to resist the temptation to ask them to catch up when they next attended. Students were encouraged to be honest when completing surveys and participating in interviews, but this may have proved awkward because they were aware that the researcher was also their tutor. It is therefore likely that their responses were, to some extent, influenced by role expectations; hence demand characteristics were at play (Orne & Whitehouse, 2000).

One means of easing the tension when being both a practitioner and a researcher is to adopt an ethical way of thinking about and conducting research, which is described as ethicality. Ethicality is influenced by axiology; that is, the researcher's values, attitudes, and beliefs as well as their professional values and conceptions of enquiry (McArdle et al., 2015). Viewing ethics as a stance rather than a set of guidelines ensured that considering the well-being of the participants in this study was not simply an item to be checked off on a list. This was particularly pertinent during Study 1 when it became evident that one of the participants had become overwhelmed by the coursework assignment.

Careful thought was given to data reporting. Given that this research was conducted where I am employed, it is possible that participants, many of whom remain at the same university for their undergraduate studies, could be identified from the vignettes or from their quotations (Fraser, 1997). Pseudonyms have been given to reduce the risk of breaching anonymity and protect ongoing relationships (Lunt & Fouché, 2010). The use of the term 'the researched university' is a further attempt to protect their identities.

4.10.2 Ethical Approval

This research was approved by the University's Ethics committee in compliance with the University's Code of Research Conduct and Research Ethics and followed the revised ethical protocols outlined by the British Educational Research Association (2018).

Prior to starting data collection, a short presentation which outlined the purpose and nature of this research was delivered within a timetabled session. All students were provided with a copy of the Information Sheet for Participants (Appendix D) as well as a Participant Consent Form (Appendix E). They were advised that while all students were expected to participate in the learning and teaching activities associated with the intervention as a normal part of their educational experience, those who wished to exercise their right to not participate in any of the data collection activities would not be adversely affected by doing so. In order to reduce the risk that participants felt obligated to participate, all documentation referred to the researcher as a doctoral student rather than a course tutor. Furthermore, students were advised that participants would not be rewarded either directly or indirectly by the researcher. Students who gave their consent to participate in the data collection activities were given the option of participating in the pre- and post-intervention JFK-Q only, though most consented to also take part in interviews, if selected.

On 25 May 2018, the General Data Protection Regulations (GDPR) came into force in the United Kingdom; participants were therefore provided with copies of the Full Privacy Notice (Appendix F) and Research Participant Privacy Notice (Appendix G). In line with GDPR (2018), data will be kept for a minimum of seven years and may be re-used in future research. All data is kept confidentially, with only the researcher and the researcher's supervisors having access to the raw data. Notwithstanding the fact that digital data are stored securely on password-protected drives, measures have been taken to safeguard data through anonymisation and pseudonymisation.

5 Intervention Design

5.1 Introduction

The purpose of this chapter is to explain how the five design principles identified within the literature reviewed in Chapter 3 influenced the design of the first iteration of the educational intervention. Design Principle 1 (*DP1*) was informed by Kienhues at el. (2016) and Sandoval et al. (2016), who proposed that providing opportunities for students to develop their understanding of subject practices through gaining disciplinary knowledge and personal experience will enable them to adopt adaptive beliefs. Hence *DP1* is to provide students with opportunities to develop their disciplinary knowledge and domain-specific beliefs. The Cognitive Affective Engagement Model (List and Alexander, 2017) was influential when devising Design Principle 2 (DP2), which is to allow students to choose topics which they are interested in. Design Principle 3 (*DP3*), which is to teach students how to construct arguments, draws upon research conducted by Nussbaum et al. (2008) who concluded that students should be taught about the components of arguments and that Toulmin's (1958) model of practical arguments is a useful rhetorical framework for doing so. Students also need to be taught how to read and engage with multiple documents (Braasch et al., 2013). Consequently, Design Principle 4 (*DP4*) was to equip students with the cognitive habits needed to engage with multiple documents. Both *DP4* and Design Principle 5 (*DP5*) drew upon the two-step validation model of Richter and Maier (2017), which indicates that elaborative processing of discrepant information is resource intensive. Students will only allocate cognitive resources if in a favourable motivational state; such a state can be promoted by providing students with positive feedback. Provision of timely feedback is also thought to be important when teaching students about the components of arguments (von der Mühlen et al., 2019). Hence *DP5* is to provide positive, timely feedback which motivates students to engage in effortful processing of belief-inconsistent information.

This intervention aims to promote the adoption of adaptive justification beliefs, and as such, this chapter begins by briefly discussing two broader issues associated with epistemic cognition interventions. It then focuses on the design

and implementation of the educational intervention. The three workshops that featured in the first iteration of the intervention are described, with attention being firstly drawn to the design principles that underpinned them. Session plans describe the teaching, learning and formative feedback activities that took place during each workshop and explain how these activities aimed to address the design principles. The final section of the chapter explains how the individual tutorial provided a further opportunity to provide timely, positive feedback. It also describes the reflection and action planning activities that students engaged in after this tutorial.

5.2 Implementing an Epistemic Cognition Intervention

Attention was paid to the recommendations made about the setting and duration of epistemic cognition interventions. With regards to the setting, Kienhues et al. (2016) specified that interventions which aim to bring about changes to justification beliefs should be situated in authentic classroom settings. The teaching, learning, assessment, and feedback activities of the intervention were therefore integrated into Semester 2 the Studying Science module, which all the researched students complete during the Science FY course.

There was a lack of consensus within the literature about the optimal duration of interventions. A meta-analysis of epistemic cognition interventions conducted by Cartiff et al. (2021) found that single session, week-long and month-long interventions had statistically significant effect sizes whereas longer-term interventions did not. These findings need to be treated with caution, however, as many of the previous studies used measures of conceptual knowledge to assess the effectiveness of the intervention. The level of specificity at which conceptual knowledge was measured moderated the effectiveness of the intervention. Interventions that used general measures of academic achievement, such as grade point averages, were found to have lower effect sizes than those that measured academic achievement at the topic level. Attention was therefore paid instead to Muis and Duffy (2013), who found that justification beliefs may take longer to change than other aspects of epistemic cognition. Consequently, the intervention activities were scheduled to take place over two to three months, beginning in early

February and ending in late March or April; the duration varied due to the timings of the spring break in the two academic years when the studies took place.

5.3 Semester 2 workshops

The three workshops are described below by identifying the design principles associated with each session. The implementation of these principles was influenced by my pedagogical philosophy, hence the teaching, learning and formative feedback activities sought to scaffold the development of effective cognitive habits. These workshops were scheduled over a three-week period beginning at the start of February and ending towards the end of the month.

5.3.1 Workshop 1 – Controversial Socio-Scientific Issues

Workshop 1 (WS1) was underpinned by Design Principle 1 (*DP1*). This was concerned with fostering adaptive domain-specific beliefs through enabling students to gain knowledge and experience of a discipline, including how research is conducted. Beliefs in justification by multiple sources are thought to be the most adaptive for science students. When engaging with multiple sources, science students need to know that scientists can hold differing perspectives about scientific issues and there may be an imbalance between conflicting positions. Students also need to know that scientists typically expect evidence to be empirical and subject to peer-review. Furthermore, the research conducted by scientists to provide evidence in support of claims often has societal impacts. In WS1, this design principle was implemented by asking students to engage with the controversy associated with the mumps, measles, and rubella (MMR) vaccine. In the late 1990s, this vaccine was incorrectly implicated as a cause of autism. The flawed study which gave rise to this unsubstantiated claim was later retracted, but not before there had been extensive media interest, which often gave equal weight to personal accounts as to scientific evidence. Due to the media coverage and the subsequent political involvement, many parents did not consent to their children being given the MMR vaccine. Hence there were significant societal impacts. As was noted in Chapter 2, public opinions about the safety of mass vaccination

programmes at the time this research was conducted were not as extreme as they became during the Covid-19 pandemic.

WS1 was also underpinned by Design Principle 4 (*DP4*), which was concerned with equipping students with the cognitive habits needed to engage with multiple documents. The documents model framework emphasises the importance of attending to source and content information. During WS1, students were guided through a process of critical evaluation which required them to consider the authors' credentials and motivations, as well as asking them to attend to the conclusions reached by the researchers and whether these were supported by the evidence. *DP4* also draws on the two-step validation model, which recognises that reading multiple documents is cognitively demanding. One means of motivating students to engage in the effortful processing of information is through the provision of timely, positive feedback; this is Design Principle 5 (*DP5*). WS1 provided students with the opportunity to gain feedback from their peers and the tutor about their initial attempts at critical evaluation of source and content information.

Session Plan for Workshop 1

<u>Introduction to the controversial issue</u> (10 minutes)

The session began with the tutor explaining that the purpose of this session was to consider a controversial socio-scientific issue: the combined MMR vaccine and its alleged links with autism. Six images of newspaper headlines, all of which referred to MMR, were displayed. Two of the headlines made specific reference to autism. The tutor explained that the MMR vaccine was incorrectly linked with an increase in cases of autism. Students were also made aware that almost all the scientists associated with the original study no longer believe there to be a link between the MMR vaccine and autism. This exemplified how opinions change in the light of new evidence. This relates to *DP1* because it demonstrates scientific knowledge is not immutable.

<u>Activity 1: Feedback from the pre-session preparation</u> (10 minutes)

The tutor reminded students that they had been given access to two documents prior to the workshop: an article entitled 'MMR and the media – a historical perspective', which was published in the News and Features section of the Society for Applied Microbiology website, and a research paper by Lewis and Speers (2003) which was published in Nature Reviews and titled 'Misleading media reporting? The MMR story'. Both documents describe a press conference given by the lead author of the now-retracted MMR study. Students were informed that during this press conference, the lead author advocated the use of single vaccines for measles, mumps, and rubella instead of the triple MMR vaccine during this press conference, even though there was no evidence in support of single vaccines in The Lancet paper. This prompted discussion about the importance of assessing whether the conclusions match the research findings, one of the cognitive habits *DP4* is seeking to develop.

The pre-session materials also described the media coverage of the controversy, noting that news reports often featured members of the public describing personal experiences. There was a brief discussion about how the public are more likely to engage with anecdotal accounts whereas scientists typically place their trust in statistically analysed data; what constitutes scientific evidence is an aspect of *DP1*. Students were also asked to consider how so-called balanced reporting can create a false impression that conflicting perspectives enjoy equal support. This relates to *DP4* because it requires students consider discrepant information.

<u>Activity 2: Critically evaluating scientific evidence</u> (25 minutes)

This activity began with students being advised that they going to read the original paper and critically evaluate the evidence presented. The tutor then displayed a PowerPoint slide which listed eight questions to provide a framework for critical evaluation, another aspect of *DP4*. These questions were inspired by work done by Helen Webster at Newcastle University to support undergraduate students in reading and critically analysing journal articles. The questions were:

1. Why were the researchers interested in this research question?

2. Do the researchers have the authority to be conducting this research?
3. How representative was this sample?
4. Are there sufficient details of how the data were collected for the study to be replicated?
5. Did the researchers choose the correct statistical test to analyse the data?
6. Can the findings be interpreted in a different way?
7. How did the findings of this study relate to previous knowledge?
8. Did this study create new knowledge?

Students were asked if they understood what these questions meant and knew what they needed to do when reading the journal article. After addressing any questions, the tutor then directed to the students to a hyperlink in the virtual learning environment to access the journal article electronically. Those without internet-enabled devices were provided with paper copies and advised to partner with students with online access when answering question 2 so that they were able to search for further information about the researchers. The tutor circulated to monitor progress, check understanding, and address any queries that arose.

Activity 3: Discussing and giving feedback about critical evaluation (40 minutes) Students working alone were asked form pairs or threes. The class was informed that they should spend 15 minutes discussing their responses to the eight questions. This activity sought to reinforce the cognitive habits associated with *DP4*. The tutor circulated during this time. After 15 minutes, students were asked to join with another pair / three to form a group of four to six students. Each group was informed which question they would be asked to give feedback on. They were given ten minutes to prepare their response and then invited, in number order, to share their feedback with the rest of the class. Each group received praise for the positive aspects of their contributions (*DP5*). Where necessary, suggestions for improvements were offered.

Activity 4: The social implications of scientific research (10 minutes)

The final part of the session provided an opportunity for students to consider the social implications of the MMR controversy. A histogram showing the percentage uptake of the MMR vaccine in England from 1997 to 2012, which had been sourced from the NHS website, was displayed. This histogram was overlaid with a line graph showing the number of confirmed cases of measles, mumps, and rubella over the same timeframe. Students were asked to describe the relationship between the uptake of the MMR vaccination and the number of cases of measles over time. The data were interpreted in different ways, with some students concluding that the decrease in the update of the MMR vaccine led to an increase in the number of measles cases but others being less convinced about a causal relationship between the two variables. As such, it illustrated how perspectives differ among scientists, an aspect of *DP1*.

5.3.2 Workshop 2 – Constructing Arguments

Workshop 2 (WS2) was grounded on the assumption that students often lack experience of engaging in argumentation, particularly in science. Nussbaum et al. (2008) advocated the use of Toulmin's (1958) model of practical arguments as a means of teaching students about the components of arguments. They also indicated that students also need to know how these components relate to each other. This research underpinned Design Principle 3 and its implementation. Along with Duschl and Osborne (2002), Nussbaum et al. recommended that students be given opportunities to analyse arguments written by others, ideally by engaging in dialogue. Consequently, students were firstly introduced to Toulmin's model before being given edited extracts of arguments written by previous Science FY students and directed to identify the different components within these arguments. These extracts featured a range of socio-scientific issues which previous Science FY students were interested in. Having seen examples of topics that previous students were interested in, the current students were encouraged to begin thinking about their own interests. Thus, WS2 began to implement Design Principle 2 (DP2), which was concerned with fostering affective engagement. When combined with effective cognitive habits, high levels of affective engagement are thought to

promote the adoption of a critical analytic stance, which is adaptive for multiple document comprehension (List & Alexander, 2017).

WS2 also provided an opportunity for students to develop their competence in constructing their own argument. Here the issue of vaccine safety, which had been introduced in WS1, provided the context. The process was broken down into three stages: providing and evaluating evidence in support of a claim, considering a counterclaim, and providing a rebuttal. As students proceeded through each of these stages, they were asked to consider the credibility of different types and sources of evidence. As such, WS2 also partly addressed aspects of *DP4*, the cognitive habits needed to engage with multiple documents. This workshop concluded with a reminder of the main components of an argument and the importance of engaging in critical evaluation, thus further reinforcing *DP3* and *DP4*.

Session Plan for Workshop 2

<u>Introduction to the structure of arguments</u> (10-15 minutes)

WS2 began by introducing students to the argument model set out by Toulmin in 1958. Attention was drawn to two components of this framework, the claim, and the data, which is now more commonly referred to as evidence. The text read: 'When we make an assertion, we are committing to the claim which any assertion necessarily involves. In order to justify / defend the claim, we need to provide facts which offer support; these are known as the data'. The tutor then explained that the way in which the data are related to the claim is known as the warrant and gave three simple statements to illustrate how warrants can be structured in order to relate the evidence (D) to the claim (C). These statements were:

- If D, then C;
- Data such as D allow one to draw conclusions, or make claims, such as C;
- Given data D, one may take it that C.

Students were then introduced the idea that warrants may feature qualifiers. The text read: 'Some warrants permit claims to be made, and accepted, unequivocally whereas others are subject to conditions, exceptions or qualifications. In such

cases, the warrant will include a qualifier (Q).' Students were shown an example from Toulmin's book to illustrate how a qualifier, claim, evidence (labelled as data), warrant, and rebuttal can be combined to construct an argument.

This example was:

- Harry was born in Bermuda (D)
- Since a man born in Bermuda will generally be a British subject (W)
- So, presumably (Q), Harry is a British subject (C)
- Unless he has become a naturalised American (R)

Activity 1: Assessing the arguments of others (30 minutes)

The tutor explained that the purpose of this activity was for students to become familiar with the components of an argument and how they relate to each other. As such, this activity was expected to play an important role in implementing *DP3*. Students were guided to view three extracts from previous assignments that had been edited for teaching purposes. Extract 1 was taken from an assignment about whether exposure to acrylamide causes cancer; extract 2 was from an assignment which asked whether there is a relationship between air pollution and type 2 diabetes and extract 3 was from an assignment about whether genetically manipulated foods have health benefits. Attention was drawn to the titles of these extracts to explain that the students who wrote the assignments were all interested in how environmental factors impact upon human health, but each was able to find a specific topic which matched their personal interests. Students were reminded that they would be expected to choose a controversial socio-scientific issue which interested them (DP2). They were then directed to read each extract and assess whether the paragraph included a claim, evidence in support of the claim and a warrant (*DP3*). They were also asked to look for critical evaluation of evidence (*DP4*). Students were guided to begin reading on their own before discussing with a partner, with these discussions focusing on how the structure of the argument could have been improved.

The tutor allowed ten minutes for students to read through the extracts individually before reminding them to form pairs to discuss their thoughts and suggest improvements. The tutor circulated during the pair discussions to check progress and respond to any queries. The tutor then led a class discussion, with students being invited to identify components of arguments within the extracts and offer suggestions about how the arguments featured in the extracts could be improved. The tutor praised students who correctly identified the claim, evidence, or warrant, or provided constructive comments (*DP5*).

Activity 2: Making an argument about the safety of the MMR vaccine (50 minutes)
The second activity in this session provided students with an opportunity to work collaboratively in a group of 3 or 4 to complete a three-part activity:

1. Providing and evaluating evidence in support of claims

The first part of this task required students to skim-read a meta-analysis conducted by Taylor et al. (2014) to find evidence in support of the claim that vaccines are safe. Finding evidence in support of claims is central to constructing arguments (*DP3*). The second part asked them to read an article written by Lewis and Speers (2003) which asserts that the MMR case was one of the earliest examples of where anecdotal evidence provided by members of the public was being presented by the media as being equal to evidence gathered through scientific means. The students were then asked to consider whether anecdotal evidence is equivalent to evidence gathered by scientific research. Being able to evaluate the credibility of evidence is an aspect of *DP4*.

2. Considering counterclaims

Students were directed to view four webpages making counterclaims. These webpages were hosted on the websites of The Healthy Home Economist, The People's Chemist, Stop Mandatory Vaccination, and the Children's Medical Safety Research Institute. They were directed to pay attention to the type of evidence provided in support of the claim that vaccines are not safe (*DP3*) and assess how credible they thought the evidence was (*DP4*).

3. Providing rebuttals to counterclaims

Students were directed to four different webpages to find rebuttals to the counterclaims (*DP3*). These webpages were hosted on the websites of the World Health Organisation, the Centers for Disease Control and Prevention, The Conversation, and Taylor and Francis online. They were asked to compare these websites to those which made counterclaims by considering the currency, authority, reliability, and purpose of each website (*DP4*).

Students were given 30 minutes to complete these three tasks. They were advised that they could allocate different parts of the task to individual members of their group but would need to allow 5-10 minutes to share their ideas before the end of the allotted time. The tutor circulated and provided regular time checks. The tutor then directed students to form groups of 6-8 students. Each group was informed which of the three activities they would give feedback about, given 10 minutes to prepare, and then invited to share their thoughts with the rest of the class. The tutor praised each group for their contributions (*DP5*), and where necessary, provided constructive feedback.

Take home messages (5 minutes)

The tutor concluded the session by restating the components of an argument (*DP3*) and reinforcing the importance of evaluating the evidence and its source (*DP4*).

5.3.3 Workshop 3 - Reading and Evaluating Sources

Workshop 3 (WS3) was primarily concerned with consolidating *DP4*. Hence the teaching and learning activities associated with this session aimed to ensure that students were equipped with the cognitive habits needed to engage with multiple documents. WS3 had a particular focus on the content of a document, and as such sought to provide students with strategies for reading complex and detailed information effectively and efficiently. The article (McGregor et al., 1994) used within WS3 was selected on the basis that the tutor was very familiar with how the research had been conducted because they had collected the data during their

undergraduate studies and had co-authored the paper. This enabled the tutor to explain the rationale behind the sampling approach, expand upon the data collection and analysis methods, and explain the findings.

Session Plan for Workshop 3

Introduction to the activity (5 minutes)

The tutor began this session by explaining that the purpose of the session was for students to develop strategies for reading journal articles effectively and efficiently (*DP4*).

Activity 1: Using figures, article titles, and abstracts to understand what the research was about (20 minutes)

This activity began with the tutor displaying a figure from a scientific paper and asking students to consider what the journal article was about. Students were invited to share their ideas with their peers. The tutor then added the figure legend and asked the students to think again about what the journal article was about. The purpose of this activity was to illustrate how figures and their legends can provide important and useful information about the findings of research. The students were then shown the title of the journal article and a brief explanation was given about the research. This was quickly followed by the tutor talking about the value of the abstract as a means of understanding the key ideas that will be discussed within a research paper. Students were then asked to read the abstract and identify the research question, sample, data collection and analysis methods, key findings, and the researchers' conclusions. They were given five minutes to complete this task before being asked them to share their ideas. The tutor gave praise (*DP5*) and offered verbal feedback.

Activity 2: Reading an extract from a journal article to find information (30 minutes)

Students were asked to organise themselves into groups of three or four. The tutor then explained that the purpose of this activity was for them to understand what

information is found in each section (Introduction, Methods, Results, Discussion) of a journal article. In doing so, they would learn where to look within an article if they needed to find further details about the research. For example, if seeking details about data analysis, students would learn these could be found in the Results section. Thus, this activity was seeking to develop the cognitive habits needed for dealing with multiple documents (*DP4*).

Each group was given a set of four extracts, with each member of the group reading one or two of the extracts and deciding which section of the journal article it was taken from. Students were given ten minutes to read the extracts. The group were then given ten minutes to reach consensus about who had read the Introduction, Method, Results, and Discussion. The tutor circulated to monitor progress, check understanding, and address any queries that arose.

After 20 minutes, the tutor asked each group in turn to indicate which extract they thought was from the Introduction and justify their decision. This was then repeated with the Method, Results, and Discussion. The students were given positive verbal feedback (*DP5*).

Activity 3: Finding key details from a journal article (15 minutes)

During activity 3, students were directed to read the full journal article and record any further details about the research question, participants, data collection and analysis methods, results, and conclusions by looking at the relevant sections of the paper. They were also asked to consider what the limitations of the study were. The tutor circulated to monitor progress, check understanding, and address any queries that arose. The tutor then emphasised how they had been provided with a framework for recording key details from a document which could be used within their argument to provide evidence, as counterclaims or rebuttals, or to inform the warrant.

Activity 4: Critical evaluation of published research (30 minutes)

This activity was a group task, with students working together to answer seven questions inspired by the work of Helen Webster that required the students to engage in critical evaluation:

1. Why were the researchers interested in this research question?
2. How representative were this sample?
3. Are there sufficient details of how the data were collected for the study to be replicated?
4. Did the researchers choose the correct statistical test to analyse the data?
5. Can the results be interpreted in a different way?
6. How did the findings of this study relate to previous knowledge?
7. Did this study create new knowledge?

The tutor spent a few minutes with each group, talking with them about whichever question they were answering at the time and providing positive feedback about their contributions to the discussion.

Summary (5 minutes)

The tutor concluded this session by summarising the strategies for reading journal articles effectively and efficiently. Students were reminded of how paying attention to the title of the article, the abstract, and figures within the text could enable them to get a good grasp of the content. They were then reminded of the value of reading with purpose, thinking about where they would find further details within the different sections of the article. They were also encouraged to engage in critical evaluation as they read.

5.4 Formative Feedback, Reflection, and Action Planning Activities

Design Principle 5 emphasised the value of providing students with timely, positive feedback to motivate them to continue engaging in the effortful processing of belief-inconsistent information. To provide students with the opportunity to receive and reflect on feedback, the intervention featured an individual tutorial and a piece of reflective writing, the Reflective Commentary assignment. This section

begins by describing what students were expected to do in preparation for their tutorial and ends with a description of the Reflective Commentary assignment.

5.4.1 Individual Tutorials

The individual tutorials were scheduled during March, beginning two weeks after WS3 had taken place. In preparation for their tutorial, students were expected to have selected a controversial socio-scientific issue, selected some of the sources they would use when constructing their argument, and engaged in critical evaluation of these sources. Each student was individually scheduled for a 15-minute tutorial. These tutorials typically began with the student being asked to identify their controversial issue and explain why they had chosen this issue.

The tutor then asked the student to discuss the sources they had selected; based on prior experience, the expectation was that most students would have only selected sources that agreed with the position they were adopting and that the tutor would therefore need to advise them to also search for alternative perspectives. The tutor also expected to provide advice about how to critically evaluate source and content information. Students were also encouraged to seek clarification about the components of an argument and about how to construct an argument from these components.

The tutorials were audio-recorded, with the resulting audio files being uploaded to the virtual learning environment so that students could listen back to the conversation when preparing their 'Reflective Commentary' assignment.

5.4.2 Reflective Commentary Assignment

Having received timely feedback during the individual tutorial, the Reflective Commentary assignment sought to ensure that students reflected and acted on this feedback when constructing their arguments. The submission deadline for the Reflective Commentary assignment was the same date as for the Making an Argument assignment. Due to the scheduling of the individual tutorials over a three-week period in March, students had between four and six weeks to complete their Reflective Commentary.

5.5 Summary

The intervention design was guided by the five Design Principles, which were implemented through the teaching, learning, assessment, and feedback activities associated with three workshops, a tutorial, and a piece of reflective writing.

DP1, provide students with opportunities to develop their disciplinary knowledge and domain-specific beliefs, underpinned the design of WS1, where students began to learn that although scientists expect claims to be supported by empirical evidence, they often interpret the findings differently. The MMR controversy provided a powerful example of how different scientists within the same research team can hold very different perspectives about the same data.

DP2, allow students to choose topics they are interested in, was firstly addressed in WS2 and subsequently during the individual tutorial. The assignments selected for WS2 provided examples of socio-scientific issues that previous students were interested in. The individual tutorials provided an opportunity for students to explain their interest in the issue they had chosen.

DP3, teach students how to construct arguments, strongly influenced the design of WS2. Students were firstly taught about the components of arguments by considering Toulmin's somewhat abstract example before viewing edited extracts of work completed by their peers. The individual tutorial provided a valuable opportunity for students who had struggled to comprehend the terms associated with an argument, such as warrant, counterclaim, and rebuttal, to ask for clarification.

DP4, equip students with the cognitive habits needed to engage with multiple documents, featured in all three workshops and was also addressed during the individual tutorials. Having asked students to critically evaluate the content of The Lancet paper that sparked the MMR controversy during WS1, the focus shifted to equipping students with the cognitive habits needed to deal with source information in WS2, with students being asked to assess the credibility of online

sources that espoused pro- and anti-vaccination views. WS3 aimed to provide students with a framework for managing information from multiple documents. The individual tutorial enabled the tutor to remind students about this framework and provide clarification about the role that critical evaluation plays when constructing arguments. In addition, the tutor was able to identify and advise students who were employing sub-optimal multiple document comprehension strategies about how to manage source and content information more effectively.

The individual tutorial and Reflective Commentary were conceptualised as a means of implementing *DP5*. The feedback, reflection, and action planning activities associated with these aspects of the intervention sought to motivate students to continue to engage in effortful processing by providing students with timely, positive feedback that would guide their future actions.

6 Study 1

6.1 Introduction

The Literature Review (Chapter 3) articulated the value of science students learning how to construct scientific arguments (Kuhn, 1993). Being able to engage in scientific argumentation enables science students to become active members of scientific communities (Rouet & Britt, 2014) and contribute to societal debates about controversial socio-scientific issues (Andriessen & Baker, 2014; Osborne et al., 2013; Simon et al., 2006). Research has shown that secondary school science classrooms provide scant opportunities for students to engage with conflicting perspectives (for example, Duschl & Osborne, 2002). Even if students have engaged in argumentation prior to starting university, they are likely to lack domain-specific beliefs about what constitutes a knowledge claim in the sciences (Green & Hood, 2013). Consequently, learning how to construct scientific arguments has become an important aspect of the Science FY curriculum at the researched university (Chapter 2).

The research reviewed in Chapter 3 drew on three areas of literature when considering the challenges that students face. Models of multiple document comprehension predict that students are likely to struggle to integrate and critically evaluate the content of multiple documents if the information conflicts with their prior knowledge and pre-existing views (Richter & Maier, 2017) or if they are not affectively engaged (List & Alexander, 2017). Argumentation research suggests that students will only be able to use the information from multiple documents to construct arguments if they have been taught about the components of arguments and how these are combined to create an effective argument (von der Mühlen et al., 2019). Justification beliefs interact with multiple document comprehension and argumentation (Bråten et al., 2014). Students who rely on personal experiences to justify why they believe something to be true attend less to the content of documents than those with beliefs in authority (Strømsø et al., 2011). Those who justify their knowledge claims by looking for corroborating evidence in multiple documents are much more likely to engage in argumentation (Bråten et al., 2014).

Hence, beliefs in justification by multiple documents are thought to be most adaptive and beliefs in personal justification least adaptive.

This reviewed research informed the design of the intervention which aimed to promote the adoption of adaptive justification beliefs (see Chapter 5 and *RQ3*: What the key design principles of an intervention designed to support students at the researched university in constructing an argument from multiple documents?). The main objective of Study 1 was to refine this intervention by asking the interviewed participants about their lived experiences of constructing an argument from multiple documents. As such, Study 1 enabled *RQ1*: What challenges do students face when constructing an argument from multiple documents? to be addressed. Study 1 participants were also asked to comment on their experiences of the designed educational intervention, which enabled *RQ4*: What are students' experiences of and outcomes from this intervention? to be partly addressed. The answers to these two research questions were used to explore the extent to which the designed intervention achieved its ambition and inform the refinements to be made before the second iteration. As Science FY students were expected to have different justification beliefs prior to the intervention, it was important to establish what these beliefs were in order to explore whether the intervention was experienced differently and resulted in different outcomes for those with different beliefs. Consequently, Study 1 also addressed *RQ2*: What are the justification beliefs of Science FY students?

6.2 Participants

The Study 1 participants were drawn from a cohort of 35 Science FY students at the researched university. 28 students consented to take part in the research, with 25 completing both the pre- and post-intervention questionnaires. 14 of these identified as female and 11 as male. The median and modal ages were both 19 years.

6.3 Findings

This section starts by describing the justification beliefs of the Study 1 cohort, clusters, and interviewees, which addresses *RQ2*. The descriptions of the

interviewees provided within the vignettes begin to address *RQ4* because they draw attention to whether DP2 was effective in allowing students to select socio-scientific issues that were of interest to them. If interviewees spoke about their prior knowledge of their chosen issue, this information is included in the vignette. Having limited content knowledge was one of four factors surfaced within the reviewed literature that informed the thematic framework developed for this research.

The next section of this chapter considers other challenges that students faced when trying to construct an argument about a controversial socio-scientific issue using multiple documents; these findings address *RQ1.* The final section describes participants' experiences of and outcomes from the educational intervention. This section ends with statistical analysis of the changes to justification beliefs, thus addressing *RQ4*.

6.3.1 Justification Beliefs of the Researched Students

At the cohort level (*N*=25), Study 1 participants had relatively weak beliefs in personal justification, moderate beliefs in authority, and relatively strong beliefs in multiple sources. There was, however, much variability between participants, with scores for beliefs in personal justification ranging from 1.00 to 6.33 and those for authority ranging from 2.00 to 9.33. The range for beliefs in multiple sources was much smaller, with the lowest value being 6.80 and the highest value being 9.60. Table 6-1 describes the justification beliefs of the Study 1 participants.

Table 6-1 *Descriptive Statistics for the Study 1 Cohort (N=25)*

Personal Justification			Authority			Multiple Sources		
Mean	Median	Range	Mean	Median	Range	Mean	Median	Range
3.56	3.33	5.33	6.15	6.33	7.33	8.05	7.80	2.80

The pre-intervention scores for the Study 1 participants were subjected to cluster analysis by Ward's method. Visual inspection of the resulting dendrogram, which can be found in Appendix H, indicated there were three clusters. These clusters differed with respect to their median scores for beliefs in personal justification and authority. Table 6-2 displays the descriptive statistics for each cluster.

Table 6-2 *Descriptive Statistics for the Pre-Intervention Study 1 Clusters*

Cluster	*n*	Personal Justification		Authority		Multiple Sources	
		Median	Range	Median	Range	Median	Range
Multiplists	10	4.67	2.33	6.25	2.33	7.70	1.80
Sceptics	6	3.17	2.67	3.83	3.67	7.70	1.80
Dogmatists	9	2.33	2.33	7.33	3.00	8.60	2.80

The names assigned to these clusters were influenced by Kuhn's (1991) developmental model of informal reasoning and Greene et al.'s (2008) epistemic and ontological cognitive developmental model. The rationale for naming each cluster is discussed further below. There are also details of the two interviewees selected from each cluster. These individuals had different magnitudes of changes to beliefs in personal justification from pre- to post- intervention: that is, for each cluster, there was one participant whose mean score in beliefs in personal justification changed the most and one participant whose mean score for personal justification beliefs changed the least.

Multiplists

The largest of the Study 1 clusters (*n* = 10) was the most challenging to name because the participants within this cluster did not fully align with either of the models of epistemic cognition adopted in this research. There was partial alignment with Greene et al.'s (2008) realists because the median value for personal justification was relatively high compared to the cohort. But the participants in this cluster did not have particularly strong beliefs in authority; indeed, the median

value was very close to that of the cohort. Moreover, Greene et al. proposed that realists would have strong beliefs in simple and certain knowledge. This description was not true of the two participants, Mollie and Matthew, selected from this cluster. Although Mollie appeared to be seeking certainty when reading multiple sources, this was not because she had beliefs in simple and certain knowledge. Indeed, she was acutely aware of the conflicting perspectives within the literature. She was better described as being a multiplist, a term drawn from Kuhn (1991), because she initially believed all viewpoints to be equally valid. This was also somewhat true of Matthew. Hence this group was named the multiplists to indicate it included students who recognise that different perspectives exist but are not yet able to evaluate the validity of these perspectives.

The pre- and post-intervention scores for the justification beliefs of Matthew and Mollie are shown in Table 6-3. Matthew was selected because he had the largest change in personal justification beliefs of those available for selection from this cluster. Mollie had the least change. Both reported decreases in their personal justification beliefs from pre- to post-intervention.

Table 6-3 *Pre- and Post-Intervention Mean Justification Beliefs of the Interviewed Study 1 Multiplists*

Interviewees	Personal Justification		Authority		Multiple Sources	
	Pre	Post	Pre	Post	Pre	Post
Matthew	4.33	3.33	5.33	6.83	7.80	10.00
Mollie	4.00	3.67	7.00	6.17	8.60	9.40

Matthew's socio-scientific issue was about whether the UK government were justified in banning neonicotinoids, which are thought to be harmful to pollinators such as bees. He became aware of this issue through seeing media reports about the European Union ban on the use of a neonicotinoids. Matthew had a long-standing personal interest in bees but had little prior knowledge of bee biology or neonicotinoids. He decided to argue that the UK government were justified in banning neonicotinoids.

Mollie was interested in the effectiveness of acupuncture in treating pain and healing injuries. She chose this issue because it was pertinent to her career aspirations. Mollie had little prior knowledge of the scientific mechanisms associated with the use of acupuncture and was hoping to find definitive explanations within the literature. Instead, she encountered conflicting scientific hypotheses and little consensus.

Sceptics

In their model, Greene et al. (2008) described those with strong beliefs in personal justification and weak beliefs in authority as being sceptics. In this research, the participants (*n*=6) in the sceptics cluster had moderate rather than strong beliefs in personal justification, with median value for the cluster being close to that of the cohort. The sceptics' beliefs in authority were weak, with the median value being the lowest of all the cohorts. The participants selected from this cluster were Stefan and Sophia, whose mean scores are shown in Table 6-4. Stefan's personal justification beliefs changed the most. His score decreased from pre- to post-intervention. Sophia was selected because her personal justification beliefs changed the least, showing a small increase from pre- to post-intervention.

Table 6-4 *Pre- and Post-Intervention Mean Justification Beliefs of the Interviewed Multiplists*

Interviewees	Personal Justification		Authority		Multiple Sources	
	Pre	Post	Pre	Post	Pre	Post
Stefan	2.67	1.67	4.33	4.83	8.00	8.40
Sophia	5.00	5.67	2.67	3.17	9.40	9.40

Stefan chose to explore the effectiveness of intermittent fasting (IF) for managing specific health conditions. He had first-hand experience of IF as a means of managing his own medical condition. His existing knowledge was based on articles written for the wider population and he was keen to better understand research about IF, particularly in relation to the effects on gene expression. He also wanted

to know whether claims made in the media about the health benefits of IF were justified.

Sophia differed from other participants in that she had selected an issue in which she held strong personal beliefs; she was opposed to research using embryonic stem cells. Although she recognised that being passionate about this this issue put her at risk of being biased, she said that she was keen to learn more about why others held views that differed from hers.

Dogmatists

This cluster was named in accordance with a term used by Greene et al. (2008), where it described students with strong beliefs in authority and weak beliefs in personal justification. The participants, Dylan and Demi, selected from this cluster (*n*=9) matched this description well because they had strong beliefs for justification by authority and weak beliefs for personal justification. Their pre- and post-intervention scores are shown in Table 6-5. Dylan's personal justification beliefs changed the most, with his mean showing a large decrease. Demi's mean score for personal justification beliefs did not change.

Table 6-5 *Pre- and Post-Intervention Mean Justification Beliefs of the Interviewed Dogmatists*

Interviewees	Personal Justification		Authority		Multiple Sources	
	Pre	Post	Pre	Post	Pre	Post
Dylan	2.33	1.00	7.00	7.50	9.60	8.20
Demi	3.00	3.00	7.33	7.33	8.60	9.00

Dylan chose the issue of whether recreational cannabis use should be decriminalised because it was an issue which could be explored from differing perspectives. For example, there had been recent media reports about research into anti-cancer and analgesic effects of cannabis. Dylan argued in support of the position that matched his own, which was that recreational cannabis use should not

be decriminalised, though he did not see himself as having strongly held views and instead described himself as being impartial.

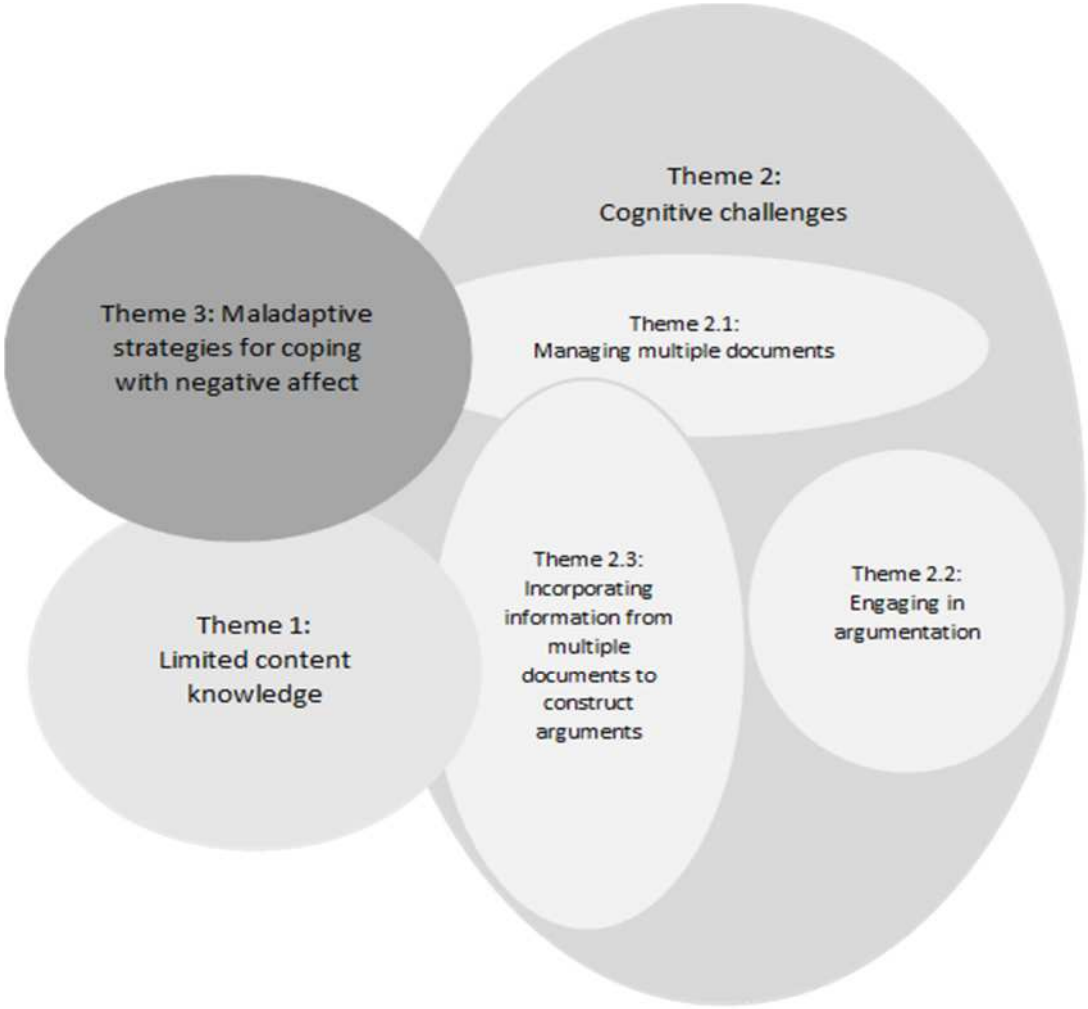
Demi selected her issue after reading a media article about research suggesting that consumption of what are termed 'ultra-processed foods' significantly increases the risk of cancer. Her choice of issue reflected her career interests, though she justified her decision by saying that she was interested in how media reporting of food issues impacts upon the dietary choices made by members of the public.

6.3.2 Challenges Faced by the Researched Students When Constructing Arguments From Multiple Documents

Chapter 4 explained how analysis by the Framework Approach (Ritchie & Spencer, 1994) enabled the development and refinement of the thematic framework for *RQ1*. The final thematic framework saw three of the themes described separately within the initial thematic framework, Managing multiple documents, Engaging in argumentation, and Incorporating information from multiple documents to construct arguments, become subthemes within Theme 2: Cognitive Challenges. There were overlaps between Managing multiple documents (Theme 2.1) and Incorporating information from multiple documents to construct arguments (Theme 2.3). These overlaps are illustrated in Figure 6-1. The final thematic framework featured two further overlapping themes: Having limited content knowledge overlapped with Having maladaptive strategies for coping with negative affect. Managing multiple documents also overlapped with Having maladaptive strategies for coping with negative affect.

Figure 6-1

The Thematic Framework for Research Question 1



The adoption of the Framework Approach (Ritchie & Spencer, 1994) enabled the experiences of the interviewees from different clusters to be compared. Table 6-6 compares the number of challenges reported by each interviewee with respect to each theme and sub-theme.

Table 6-6 *Comparing the Challenges Reported by Study 1 Interviewees*

Themes	Sub-themes	Multiplists		Sceptics		Dogmatists	
		Matthew	Mollie	Stefan	Sophia	Dylan	Demi
1. Having limited content knowledge		+		+			
2. Cognitive challenges	2.1 Managing multiple documents	+ + +	+ + + + +	+ +	+ +	+ +	+ + +
	2.2 Engaging in argumentation	+ + +	+ +	+ + +	+	+	+ +
	2.3 Incorporating information from multiple documents to construct arguments	+	+		+	+	+
3. Having maladaptive strategies for dealing with negative affect		+	+		+		+

Note: The more plus signs, the more of the challenges described by the theme or sub-themes were experienced by the interviewees. Empty boxes indicate that the interviewee did not report facing any of the challenges associated with that theme or sub-theme.

Theme 1: Limited Content Knowledge

Participants differed in their pre-existing knowledge of their chosen socio-scientific issue. Difficulties faced by those with limited content knowledge were described within the first theme of the thematic framework developed for this research. The accounts provided by Matthew and Stefan indicate that these difficulties arose when students who lacked pre-existing knowledge tried to read journal articles. Doing so created two challenges: they were faced with unfamiliar subject-specific terms, and they were presented with statistical analysis of data. In relation to the first of these, Matthew explained that he did not always understand the "complex words" used within research articles and Stefan acknowledged that there were "words I don't understand". When discussing the second of these challenges, Matthew said: "I don't actually understand the data they're showing me with the statistics" and Stefan described himself as becoming frustrated by not knowing how to interpret data, saying,

> "When you get to the results section and they are just quoting mean standard deviation and then other random statistic that they've found and it's like I don't really know any of this and it's just making me frustrated".

The experiences of Matthew and Stefan suggest that having limited content knowledge can trigger negative affect if students attempt to read sources written for those with greater disciplinary experience. Although Stefan did not report having adopted any maladaptive coping strategies in response to having experienced negative affect, Matthew did. The role that negative affect can play in the adoption of maladaptive coping strategies is further explored within Theme 3. Having limited content knowledge also overlapped with Theme 2.3, the challenges of incorporating information from multiple documents to construct arguments.

Theme 2: Cognitive Challenges

The second theme developed within the thematic framework had a much broader scope than Theme 1, encompassing three sub-themes: Managing Multiple

Documents, Engaging in Argumentation, and Incorporating Information from Multiple Documents when Constructing Arguments.

Managing Multiple Documents

The Managing Multiple Documents sub-theme included five challenges that students faced when trying to manage multiple documents. These five challenges were: difficulties in finding sources, problems of employing maladaptive search strategies, negative consequences of reading more deeply, challenges of synthesising information and difficulties associated with engaging in critical evaluation.

As indicated in Table 6-6, all interviewees experienced at least two of these challenges, with finding sources and engaging in critical evaluation being the most common. The reported experiences of one of the multiplists, Mollie, span all five aspects, hence she features prominently within this section. Time-management cut across four of the five challenges, with participants often exclaiming how they had underestimated the time needed to find, read, synthesise, or critically evaluate sources.

Interviewees from all three clusters described finding sources to be challenging. Even when source information was provided, such as in the References section of a journal article, participants reported that they were not always able to find sources quickly. The time-consuming nature of finding sources was sometimes exasperating, with Stefan remarking there were some sources "you'll probably never be able to find for three weeks". Although not particularly true of Stefan, experiencing negative affect when trying to find sources appears to have led some of the interviewees to adopt maladaptive coping strategies.

Employing maladaptive search strategies proved problematic for some of the interviewees. For example, both Demi and Mollie initially tried to develop their knowledge about their chosen socio-scientific issues using journal articles rather than review articles or textbooks. Demi soon realised that she was going to be confronted with specific details, and this was not what she was looking for,

remarking “There was so much about *the cell cycle* but I was trying to find (*more general information*)”. Demi sought advice about how to proceed during her individual tutorial. This differed from Mollie, who persisted in trying to gain an overview of the issue by continuing to read journal articles despite having spoken about how she was “overwhelmed” by the number of sources she had already read. When writing her reflective commentary, Mollie identified one of the causes of her maladaptive search strategy when she wrote, “Due to a lack of clarity whether my argument should focus on a specific or general musculoskeletal condition, I noticed my research (*sic*) strategy was too broad”. Adopting an overly broad search strategy may be particularly problematic when there is a lack of consensus in the literature. In Mollie’s case, it appears to have caused her to become uncertain about which position to adopt when making her argument:

> “There were a lot of disagreements to different things that’s why I remember coming to the meeting and saying I don’t know whether I’m for or against it … I really could go either way because there’s just so many conflicting opinions”.

Having experienced negative affect, Mollie procrastinated, illustrating the overlap between Managing Multiple Documents sub-theme and the Maladaptive Strategies for Coping with Negative Affect theme.

The Managing Multiple Documents sub-theme of the thematic framework developed within this research was informed by the models of multiple document comprehension discussed within the literature review (Chapter 3). One of these models was the Document Models Framework (Britt et al., 1999; Perfetti, et al., 1999), which describe the behaviours of skilled readers. The participants in this research were not skilled readers. As such, their experiences illuminate the challenges of constructing a coherent mental representation of multiple documents. When faced with these challenges, some participants responded by adopting surface approaches to reading. For example, Matthew and Dylan both spoke of skim-reading or highlighting. Those who attempted to adopt deeper approaches sometimes faced negative consequences. For example, Mollie talked of

how she had become so engrossed in the content that she forgot to critically assess what she was reading when she said: “I was just reading so many different things and just got so kind of deep into all of the knowledge I was reading about I forgot about that kind of stuff”.

Mollie was not alone in having experienced difficulties engaging in critical evaluation. Many of the other interviewees also expressed having faced challenges when trying to critically evaluate their sources. In some cases, this was because they were unsure about how to do so. For example, when asked if there was anything that had caused her to question the quality of her sources, Sophia replied: “I don’t know really I Googled all the people and they all seem like legitimate people I don’t really know”.

Others were reluctant to engage in critical evaluation. Matthew revealed he was afraid of having to reject sources if critical evaluation revealed flaws:

> “I’d think to myself right let’s not be afraid let’s analyse this one and then if that one’s analysed really well and you think it’s bad then don’t be afraid to ... just put it to the side and try and find another one”.

Matthew’s comments also indicate how time-management often hindered critical evaluation. Having found it difficult to find good sources, Matthew was reluctant to reject those he had found because this meant he would need to invest more time in finding replacements. Three other interviewees also reported limited engagement in critical evaluation due to time-management issues. Demi commented on how she “hadn’t put as much thought into taking what I’d looked at and critically evaluated it I’d done it to an extent but yeah that was kind of a bit rushed slightly”. Likewise, Mollie reflected on how she “didn’t critically evaluate due to poor time management underestimating how long it would take” and Stefan wrote that “more effective time management on my part would have allowed in-depth analysis of the evidence”. It therefore appears that most participants found this aspect challenging because they had underestimated the time needed to critically evaluate their sources. This is probably because they lacked experience of working with multiple documents.

Synthesising information from multiple documents also proved to be challenging. The interviewed multiplists featured prominently within this theme. For instance, Mollie said she found synthesis of sources to be "hard and just took such a long time". These remarks suggest that unrealistic expectations about the time needed for elaborative processing may have added to the difficulties she faced. Moreover, Mollie appeared to have been seeking certainty and struggled to know how to proceed when faced with inconclusive information: "There's nothing really saying actually is does definitely increase anti-inflammatory proteins or it does definitely stimulate blah blah blah'". The other interviewed multiplist Matthew also struggled with synthesis, saying: "I find when there's a lot of things going on my mind gets cluttered and I can't see almost what I'm actually trying to say". One possible explanation for these findings is that the interviewed multiplists lacked effective corroboration strategies.

Engaging in Argumentation

Participants from all three clusters described challenges they had faced when trying to engage in argumentation. It is likely that these challenges arose because most did not have prior experience of constructing scientific arguments. For example, Mollie remarked "I've never done anything like this before". Even when participants had prior experience of engaging in argumentation in other disciplines, the context in which they were now working felt new or different to them. For instance, Demi explained that although looking for key details in a document was something that she was familiar with from having studied A level English, she lacked experience in using evidence from scientific papers:

> ''cos this is all quite new … 'cos well I did English A level but we had to read similar well not obviously very different content but like the way it was so dense and everything … so I think that's come alright to me the actual reading of it but trying get what's important that's kind of in me already but it is very different because I haven't really had to do this before".

Having prior experience of engaging in argumentation through a debate society did not prevent Matthew from sometimes being reluctant to engage in argumentation. Two factors appear to have contributed to his reticence: firstly, the other students in the debate society were typically studying the humanities or social sciences, such that he had little experience of debating with other science students, and secondly, he was unsure about challenging those who he felt were better qualified, commenting "when it comes to something like this I was like you've done all this training I'm just a student trying to find my way and I remember being really cautious".

The reported experiences of Demi and Matthew suggest that having experience of engaging in argumentation is not sufficient in enabling Science FY students to feel confident in their abilities to construct a scientific argument about a socio-scientific issue.

Although all interviewees had selected a socio-scientific issue by the time of their individual, some found reaching a decision about which issue to choose to be challenging. For instance, Demi described deciding on the issue as being her "main struggle". She explained "I was flitting between a few different ones … and I was trying to decide and it made my start quite slow". Demi's experience illustrates how cognitive challenges often interacted with time-management issues.

Staying focused on their chosen issue proved to be particularly challenging for the two interviewees whose chosen issues were strongly associated with ethical or political debates. Sophia, who was arguing against the use of embryonic stem cells in research, noted that, "it's kind of hard 'cos a lot of it feeds into other stuff so then it's here's IVF rights and then a lot of embryos from abortions are used and it's abortion rights". Dylan, whose issue was whether recreational cannabis should be decriminalised, said how he thought "it's quite easy to slip into the political side of things and draw away from the science with this it's something I'll have to watch out for". Both Sophia and Dylan were arguing in support of their pre-existing views, suggesting that those who have strong ethical or political beliefs about an issue may struggle to stay focused on the scientific aspects of socio-scientific issues.

Other interviewees struggled to align their socio-scientific issue, or their position on this issue, to the claims they wanted to make within their argument.

Stefan raised his concerns during the individual tutorial, saying that he felt like he had "picked a topic and then stubbornly tried to make it fit". Matthew also struggled to match his claims with the position he had adopted, leading him to note during the retrospective interview that:

> "my claims were too separate from the overall point I was trying to make and so when it came to linking them all together it was very hard to say to make it flow it was very much kind of we're jumping over here now and now we're jumping over here".

Matthew and Stefan both sought advice during their individual tutorials and were subsequently able to successfully align their claims with their issue.

Interviewees from all three clusters spoke of two cognitive challenges associated with the structure of an argument. One of the sceptics, Stefan, had difficulties with the abstract nature of the terms used to describe the components of an argument. He explained that he found these terms difficult to understand, saying that he "couldn't think that literally about this". The multiplists and dogmatists struggled with specific components of an argument, such as the warrant, rebuttal, and counterclaim. It appears that Mollie had struggled to grasp the role of the warrant because she thought that "your evidence had to … directly link to the claim". Matthew spoke of "having trouble with (*the*) rebuttal". He had also struggled to understand the counterclaim, as did Demi and Mollie. Despite having outlined a rebuttal and counterclaim during her individual tutorial, Mollie reported that she had never really got to grips with these components, saying this "was annoying and it frustrated me". These findings suggest that the abstract nature of the terms used to describe the components of an argument can sometimes hinder students' perceptions of how well they understand an argument's structure even when they have successfully constructed their own arguments.

<u>Incorporating information from Multiple Documents to Construct Arguments</u>

Most interviewees reported experiencing difficulties when trying to incorporate information from multiple documents to construct arguments. For Matthew, these difficulties related to his limited content knowledge (Theme 1). In particular,

Matthew's lack of understanding of statistical data left him unsure of how to refer to statistics in his argument. He was also unsure how critical evaluation, which he had been reluctant to engage in, would be incorporated into his assignment. For other participants, the challenges they had faced when trying to manage multiple documents (Theme 2.1) often resulted in participants struggling to integrate source information into their arguments. For instance, the difficulties that Mollie experienced when trying to corroborate contradictory and inconclusive information caused her to comment that she was struggling to back up her claims. She later reflected on how she had "found synthesising information and providing evidence most difficult".

In common with other cognitive challenges, interviewees reported having struggled to manage their time effectively when trying to incorporate information from multiple documents to construct arguments. For example, Sophia reflected that she "hadn't spent a huge amount of time on research at that point and knew more would be required". For some, finding more time was challenging due to competing demands. For example, Mollie said how she had "stress from other deadlines".

These findings indicate that students find it challenging to incorporate information from multiple sources when constructing arguments if they have limited content knowledge, have difficulties in managing multiple documents, or do not allocate enough time.

Theme 3: Having Maladaptive Strategies for Coping with Negative Affect

The third theme within the thematic framework developed for this research described the maladaptive strategies that participants adopted in response to experiencing negative affect. Participants often vocalised negative affective responses, such as feeling "worried", "stressed", "anxious" or "frustrated", being "cautious" or "afraid", or "not being sure". These negative affective responses were often expressed when discussing the challenges of managing multiple documents (Theme 2.1), though sometimes arose from having limited content knowledge (Theme 1).

The most common response to experiencing negative affect was the need to be reassured, with almost all participants using words or phrases during their individual tutorial that indicated they were seeking reassurance. Demi explained that needing reassurance was challenging because she was working individually, saying "When you're on your own you can keep second guess yourself and there's not really anybody else to be like yes let's do that". Other participants wanted reassurance because they were afraid of making mistakes, with Matthew saying: "I was afraid I would do it wrong". Mollie was also keen to get things right, asking question such as: "should I do it in relation to a particular body part", "would be better to do something more specific or does it really matter", and "is it valid to compare acupuncture for different parts of the body" during her individual tutorial.

The need for reassurance sometimes limited how much progress students made in the early stages of the assignment. This was particularly true of the multiplists, though Sophia, a sceptic, and Demi, a dogmatist, also spoke about having slow starts due to being unsure. Procrastination is often considered to be a time-management issue, but here it appears to be one of several maladaptive strategies adopted by students who are finding things challenging. For example, Matthew's need for reassurance had caused him to leave little time for reading his sources ahead of the tutorial, leading him to adopt a surface approach to reading:

> "I was worried too much for the tutorial and I think I four days beforehand I was like oh wow I should really crack on and do stuff and I literally just went along and highlighted everything".

Matthew later struggled to relate the information he had extracted from sources to the position he wished to adopt in relation to his chosen socio-scientific issue, suggesting that his strategy of highlighting text had not been an effective means of finding relevant content from documents.

Mollie, also from the multiplist cluster, similarly reported having procrastinated. She explained that this was because she had found the process of trying to compare and synthesise conflicting information to be challenging and time-consuming: "I was putting it off because I didn't want to go back to it 'cos I was like oh I've gotta open this now and try and cut this down so I was just putting it

off". Having reported experiencing all five of the cognitive challenges associated with managing multiple documents, it is perhaps not surprising that Mollie reported the strongest negative affect of all the interviewed participants. For example, during her individual tutorial she used the words "worried" or "worrying" on five separate occasions, as well as talking of how she was "struggling" (two occasions), felt "overwhelmed and was "freaking out". While seeking reassurance proved adaptive for Mollie, her experiences highlight how the cognitive challenges posed by managing multiple documents can potentially lead multiplists to experience strong negative affect which in turn may result in them adopting maladaptive coping strategies.

6.3.3 Participants' Experiences of And Outcomes from the Intervention Activities

Although Study 1 was primarily concerned with the challenges that students faced when constructing an argument using multiple documents, interviewees often commented on their experiences of the intervention activities. Using the Framework Approach (Ritchie & Spencer, 1994), these experiences were charted against the Design Principles (DPs). The outcomes from the intervention were measured both qualitatively and quantitatively. The qualitative findings are discussed below and the statistical analysis of the changes in justification beliefs from pre- to post-intervention can be found in Section 6.3.4. Collectively, these findings address *RQ4.*

When discussing their experiences of and outcomes from the intervention during individual tutorials, retrospective interviews, or in their written reflective commentaries, the interviewees rarely commented on their experiences of being given opportunities to develop their disciplinary knowledge and domain-specific beliefs (*DP1*). Those who did typically described the challenges they faced when trying to understand documents written for expert audiences. Consequently, the interviewees' experiences of this aspect have already been discussed in relation to *RQ1*. This section therefore considers DPs 2 to 5. Where appropriate, attention is drawn to activities that were described as being particularly helpful or unhelpful.

DP2: Allow students to choose topics they are interested in

When planning the workshops, it was intended that WS2 would provide the ideal opportunity for implementing *DP2* because students would see examples of socio-scientific issues that previous students had been interested in. The only interviewee who commented on this design principle did so with reference to WS1, not WS2. Demi said the MMR controversy had been particularly helpful in aiding her in choosing her own topic:

> "I think that was such a good example of seeing an example how this actually relates in the world 'cos we were talking about socio-scientific issues but it kind of helped me think what one that … was a good example of one and it made me helped me think about what I might do".

Although there was limited evidence about whether the workshop activities had proved effective in implementing *DP2*, five of the six interviewees reported choosing socio-scientific issues they were interested in.

DP3: Teach students how to construct arguments

This design principle was implemented during WS2 when students were taught about the components of arguments and given examples of previous students' work. There were striking differences in participants' experiences of learning about Toulmin's rhetorical framework. At one extreme, Dylan described learning the components as described by Toulmin as being the most helpful activity:

> "He labelled off all the cd, q yeah that was really really helpful like sort of looking at constructing an argument and then just breaking it down very mathematically I think that will help me going on from this actually and in sort of even just debating and stuff like that it's a very logical way to construct an argument".

Stefan's experience was very different; he identified WS2 as being the least useful aspect, saying that he "didn't really take much away from that session".

Despite having very contrasting experiences, almost all interviewees arrived at their individual tutorial having identified at least two, and in many cases three, claims within the sources they had been reading. Thus, it can be concluded that WS2 had been effective in enabling students to learn what constitutes a claim. In contrast, WS2 did not successfully teach about warrants, counterclaims, and rebuttals, with all interviewees seeking clarification about these components during their individual tutorials. Moreover, when creating their action plans within the Reflective Commentary assignment, many participants remarked that they needed to better understand these components. For example, Sophia wrote: "From my feedback I knew that the area requiring the most attention was the development of a counterclaim".

Interviewees sought clarification about at least one of the components during their conversation with the tutor because they had struggled to fully comprehend the terms used to describe the different components. To provide concrete examples of what are somewhat abstract concepts, WS2 had featured edited extracts of previous students' work. It was hoped that these examples would illustrate how the components are used to construct an argument, but interviewees' experiences of these exemplars were mixed. The multiplists, Matthew and Mollie welcomed them, with Matthew describing how he had viewed the examples as a means of learning the approach he could use when constructing his own argument. Mollie similarly found looking at other people's work to be helpful, explaining that it allowed her to understand what she needed to do: 'pulling out claim evidence warrant to see what I actually had to do'. In contrast, both the sceptics did not benefit from viewing exemplars. Sophia was unable to recall being given examples. Stefan was aware of the examples, but did not find them helpful and instead offered an alternative way of teaching students about the components of arguments:

> "Just write a small sentence yourself write your claim out on one line then a couple of lines write your evidence out couple of lines and then write the rebuttal or something and then piece them together".

It therefore appears that justification beliefs may play a role in students' experiences of and outcomes from using exemplar material when learning to construct arguments, with multiplists benefitting from seeing examples of others' work but sceptics less likely to find them helpful.

DP4: Equip students with the cognitive habits needed to engage with multiple documents

DP4 influenced the design of all three workshops. WS2 appears to have been of limited effectiveness in teaching students how to assess the credibility of sources. Those who did pay attention to source information often relied on just one or two aspects, such as checking the year of publication or noting where the article had been published. For example, Sophia wrote: "When selecting my sources I specifically selected those that were published less than 15 years ago by peer reviewed journals". Furthermore, Dylan revealed there may have been some confusion about the status of subject databases when he said:

> "What I s'pose for this I looked at with regard to quality it was things like the actual journal they were published in if they were published on a journal or if they were just published on is it PubMed 'cos I know that's something like three thousand articles every day and some of them are just absolute rubbish".

Just three interviewees spoke of checking the author's affiliations. This was usually achieved through conducting an internet search. For example, Sophia explained that: "from the onset so generally before I even read the abstract of an article I would Google the author because it would just be a waste of time if they weren't a legit person".

Some felt uncomfortable in making decisions about whether to reject a document on basis of what they viewed as being subjective assessments. For example, Dylan remarked, "There's no real objective measurement of that (*author's credibility*) it was things like did they attend a university that seems credible do they have any other papers out there". Others were, however, able to make judgements about whether to trust sources. This was particularly true when authors were

affiliated with commercial organisations. For instance, Stefan commented on how "one of the researchers was from a health food brand so I think I might want to question that".

WS3 sought to provide students with a systematic approach to reading and recording important information from journal articles. While some opted to highlight or annotate texts, those who reported using the approach outlined during the workshop, which included completing a template to record key details, said that they had found it useful. Interestingly, the three interviewees (Matthew, Stefan, and Dylan) who reported having positive experiences of using the template were those whose beliefs in personal justification showed the most change from pre- to post-intervention. Matthew explained that he had "followed a method set out by the lecturer" which he found to be a "good process for understanding articles" because it allowed him "to easily see if there were any claims to be made". Stefan had also adopted the guidance given, though he had viewed the process as providing a means of determining whether a source was relevant:

> "And it was a case as you'd gone through in that lecture looking at the abstract looking at the conclusion does it talk about this this and this that I want it to in terms of insulin resistance or does it talk about this in terms of fat reduction or whatever yes okay put it in that box no get rid of it".

Dylan was similarly strategic, reporting that he had used the abstract as a means of identifying suitable sources, though he did recognise the need to read the main body of the article when looking for information to include in his argument. The experiences of Matthew, Stefan, and Dylan suggest that learning how to extract information from sources may be associated with a decrease in the reliance on personal beliefs.

Guidance about critically evaluating research was incorporated into two workshops, with students firstly being provided with a set of questions during WS1 and later being given a template to complete when reading extracts from a journal article during WS3. Those who adopted the strategies described within the workshops reported them as being effective, with Mollie describing this as being

one of the most valuable aspects, saying, “I found the questions really helpful for critically evaluating that’s just basically how I critically evaluated so I knew what I was doing there”. Yet most remained uncertain about how to critically evaluate the content of their sources. For instance, Matthew said: “I wish I knew exactly how to critically analyse something”. Others spoke of only engaging at a superficial level, though this was sometimes because they not yet adopted deeper approaches to reading. For instance, Dylan had not read beyond the abstract at the time of his individual tutorial and subsequently recognised the need to read more deeply within his action plan:

> “I will further analyse and critically evaluate my sources. I will look into the main bodies of the studies, the methods, the results and the discussion in order to find more information. Particularly for the critical analysis I will look at the methods, concentrating on the sample size and look for any potential problems or anything else that leads the studies subject to bias”.

Stefan also felt he had not yet fully engaged with critical evaluation, though it appears he was being overly critical in his self-assessment as he had, in his own words, considered the “demographic, size and species of the sample, specific illnesses of the sample, authors credentials, and dates of publication”.

The workshops were not effective in setting expectations about the amount of time needed for engaging in critical evaluation. Consequently, some students had not scheduled enough time. For instance, Demi reflected that she “hadn’t put as much thought into taking what I’d looked at and critically evaluating it like I’d done it to an extent but yeah that was kind of a bit rushed slightly”. Conversations during individual tutorials also indicated that the workshops were not always successful in enabling students to understand how to incorporate critical evaluation into their arguments. For example, during his tutorial Matthew was advised to change the wording in the warrant if he identified flaws in the evidence being used to support a claim. In response he said, “so the point of us critically analysing it was to find a level”, indicating that he now understood that he had been asked to

engage in critical evaluation so that he could determine whether claims should be qualified.

In summary, the reported experiences of the interviewees indicate that the implementation of design principle 4 was only partly effective in equipping students with the cognitive habits needed to engage with multiple sources. Although some of those interviewed had learnt how to read and record important information from sources, almost all were yet to become competent in critically evaluating source and content information.

DP5: Provide positive, timely feedback which motivates students to engage in effortful processing of belief-inconsistent information

This design principle underpinned the individual tutorial and Reflective Commentary. Reflective accounts of the feedback given during the tutorial revealed that all interviewees felt that the tutorial had led to positive outcomes. As such, it is unlikely that justification beliefs impacted upon this aspect of the intervention.

Although all reported positive experiences of the tutorial, some had been apprehensive. For example, Mollie wrote that she "was anxious about impending criticism due to my perceived lack of preparation and clarity" ahead of the tutorial whereas "to my surprise the overall feedback was positive". Indeed, Mollie later remarked that the tutorial "pretty saved me to be honest" from heading off in the wrong direction with her argument.

All participants reported how the tutorial had drawn attention to positive aspects of their work. Some participants wrote about getting positive feedback about their developing cognitive habits. For example, Dylan referred to being given positive feedback about the types of sources he had selected, the annotations he had added to these sources, and his synthesising of information. Others indicated that the feedback had identified aspects of their assignment where there was clarity. For example, Demi described how the tutor had been pleased there was "clear direction in the planning of my project and mentioned that even before I had started writing, it is clear to picture how my final essay will be structured and read".

When describing their experiences of being given feedback, four interviewees used the word “confident” or “confidence”. This confidence was typically expressed in general terms, such as Mollie saying she felt confident about the next steps, Demi being confident she was on the right lines, and Matthew feeling more confident in how to write his assignment. Sophia was more specific, commenting on how feedback about the appropriateness of her sources had given her “confidence to cite them assertively in my assignment”. She also reported feeling more confident in “knowing how to structure her response”.

Interviewees also reported experiencing positive affect because of the feedback they were given during the tutorial. Examples include Sophia feeling “encouraged”, Mollie having a “more positive attitude”, and Matthew reporting being “motivated to continue searching for a source” he was having difficulty locating. Given these wholly positive outcomes, it can be concluded the implementation of *DP5* was the most effective aspect of the first iteration of the intervention.

6.3.4 Changes to Participants’ Justification Beliefs

The differences between pre- and post-intervention JFK-Q scores provided a quantitative outcome measure of the changes to participants’ justification beliefs at the cohort level *(N* = 25).

Table 6-7 *Descriptive Statistics for Participants Pre- and Post-Intervention Beliefs in Study 1*

Justification Belief	Pre-Intervention				Post-Intervention			
	Mean	SD	Median	Range	Mean	SD	Median	Range
Personal Justification	3.56	1.37	3.33	5.33	3.17	1.42	3.00	6.00
Authority	6.15	1.77	6.33	7.33	6.29	1.75	6.50	7.17
Multiple Sources	8.05	0.87	7.80	2.80	8.13	1.39	8.40	4.20

.

Table 6-7shows the mean and median pre- and post-intervention values for participants' beliefs in personal justification, authority, and multiple sources. As predicted, there was a decrease in the average value for beliefs in personal justification. Also as predicted, the average values for beliefs in authority increased, as did the average values for beliefs in multiple sources. Statistical analysis of the differences in median values was conducted using the one-tailed Wilcoxon matched-pairs signed-ranks tests with the Pratt's correction for tied values. The change to personal justification beliefs was approaching significance at the 0.05 level, indicating that there was a tendency for participants' personal justification beliefs to weaken. The differences to beliefs in authority and multiple sources were not statistically significant. Given the data are averaged ranks, some statisticians and previous analysis prefer parametric analysis for these data. Consequently, three paired t-tests were also conducted. These indicated that there was a significant decrease in personal justification beliefs. The results of these analyses are displayed in Table 6-8.

Table 6-8 *Statistical Analysis of Differences Between the Study 1 Pre- and Post-Intervention Scores by Justification Belief*

Justification beliefs	*W*	*p*	*t*	*p*
Personal justification	-114.0	0.0634	1.86	0.037
Authority	63.0	0.2030	.896	0.190
Multiple Sources	36.0	0.3190	.282	0.390

Further analysis of the changes to justification beliefs was conducted to assess whether participants' overall justification beliefs had become more adaptive. The values for beliefs in personal justification were firstly transformed to ensure that a decrease in personal justification beliefs gave a positive score. The values on the three sub-scales were then summed, and the pre-intervention summed value was compared to the post-intervention summed value using the Wilcoxon matched-pairs signed-ranks test for differences with the Pratt's correction for tied values. The significance level (α) was adjusted for multiple comparisons using the Bonferroni method. This was because the summed values used the same data as had been used to compare the changes to justification beliefs as measured by the sub-scales. Consequently, the *p* value could only be deemed significant if equal to or lower than 0.025. The analysis showed that there was a tendency for participants' summed justification beliefs to increase from pre- to post-intervention, $W(25) = 106.0$, $p = 0.0802$. Although not significant, this result indicated that many of the Study 1 participants had experienced positive outcomes with respect to their justification beliefs. The post-intervention JFK-Q scores were also used to sort participants into post-intervention clusters. Visual inspection of the dendrogram resulting from cluster analysis using Ward's method indicated that there were two clusters, as shown in Table 6-9. The dendrogram can be found in Appendix H.

Table 6-9 *Descriptive Statistics for the Study 1 Post-Intervention Clusters*

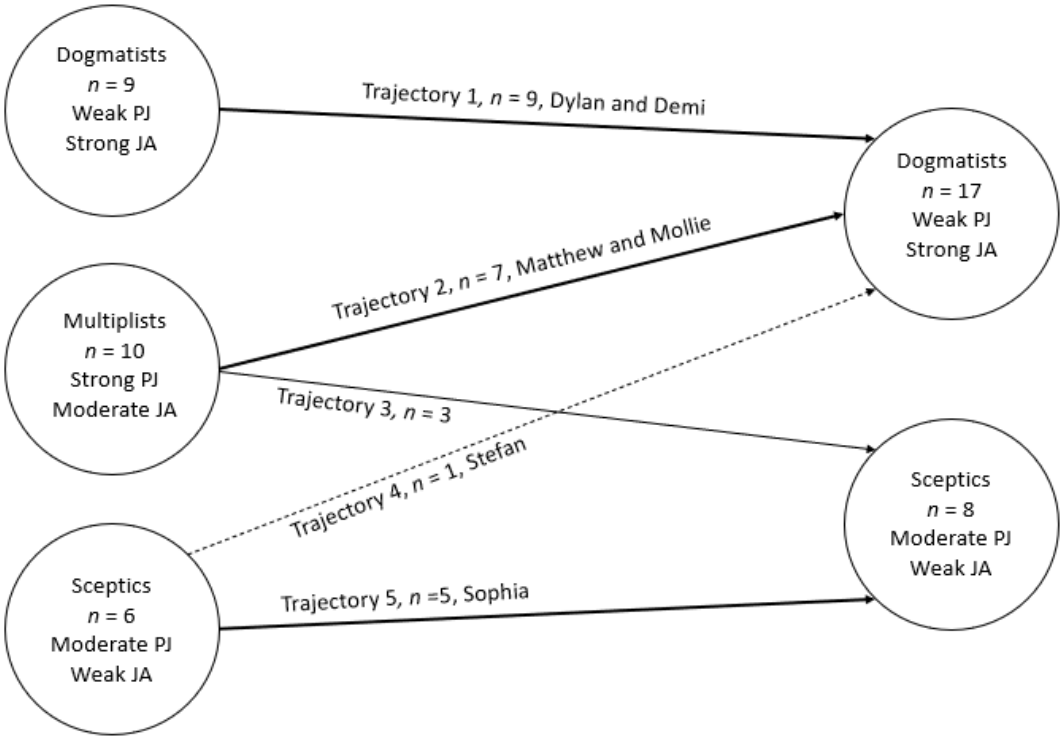
Justification Belief	Sceptics (*n*=8)		Dogmatists (*n*=17)	
	Median	Range	Median	Range
Personal Justification	4.00	4.33	2.67	4.00
Authority	4.84	2.84	7.00	4.67
Multiple Sources	7.40	3.60	9.00	4.20

The largest cluster (*n*=17) was named the dogmatists because the participants in this cluster had relatively weak beliefs in personal justification and relatively strong beliefs in authority. The post-intervention dogmatist cluster included all nine of the pre-intervention dogmatists. Of these, four reported increases, three reported decreases and two had no change to their scores for personal justification beliefs. The post-intervention dogmatist cluster also included one participant who had been in the pre-intervention sceptics cluster and seven of the pre-intervention multiplists. All these participants reported decreases in their beliefs in personal justification.

The smaller cluster (*n*=8) was named the sceptics because the median score for authority was relatively weak. Along with five of the six pre-intervention sceptics, all of whom reported increases in their personal justification beliefs, this cluster also included three multiplists. While these participants differed in the changes to their personal justification beliefs, with two reporting decreases and one increasing, all reported decreases in their beliefs in authority.

The movement of participants from their pre-intervention clusters to their post-intervention clusters was tracked. There were five trajectories, as illustrated in Figure 6-2. The interviewed participants followed four of the five trajectories. Dylan and Demi, who both remained dogmatists, followed Trajectory 1. Matthew and Mollie followed Trajectory 2 to became dogmatists. Stefan also became a dogmatist, and as such followed Trajectory 4. Sophia remained as a sceptic, following Trajectory 5.

Figure 6-2 *Trajectories from Pre- to Post-Intervention Study 1 Clusters*



Note. Changes in cluster membership before and after the intervention, with number of participants moving from each pre-intervention cluster to each post-intervention cluster. Dashed lines indicate that the proportion of students moving along that the trajectory was less than 25% of those in the pre-intervention cluster, thin solid lines indicate the proportion was between 25% and 50%, and thick solid lines indicate the proportion was at least 50%. The names indicate the interviewees who followed that trajectory.

6.3.5 Summary and Implications for the Design of Study 2

Participants in the Study 1 cohort held differing initial justification beliefs (*RQ2*). Those within the largest cluster, the multiplists, contrasted to the second largest cluster, the dogmatists, with respect to their beliefs in personal justification. Such beliefs were relatively strong in the multiplists and relatively weak in the dogmatists. The smallest cluster, the sceptics, differed from the dogmatists with

respect to their beliefs in authority, with the sceptics having relatively weak beliefs in authority and the dogmatists having relatively strong beliefs in authority. Most participants reported changes to their justification beliefs. Personal justification beliefs were expected to become weaker and statistical analysis supported this prediction. This was true for just over half (13 of 25 participants) of the Study 1 cohort, including the two interviewed multiplists, one of the interviewed dogmatists and one of the interviewed sceptics. Of the remaining participants, ten had stronger beliefs in personal justification after the intervention, including one of the interviewed dogmatists and one of the interviewed sceptics.

The challenges that the interviewees faced when trying to construct an argument about a socio-scientific issue from multiple sources (*RQ1*) were identified during Study 1. One of the key findings was that adopting maladaptive search strategies when looking for sources often leads to negative affect, which may lead to the adoption of maladaptive coping strategies. Some participants became annoyed or exasperated because they found themselves struggling to understand the complex terminology and unfamiliar concepts they encountered when attempting to use journal articles to develop their disciplinary knowledge. Others became overwhelmed by the range of perspectives expressed across the different sources, leaving them uncertain about which position to adopt within their own argument. Although participants often sought reassurance and advice during their individual tutorial, many had already adopted maladaptive coping strategies, such as procrastinating. These findings indicated that students needed to be provided with clearer guidance about the sources they should use to develop their disciplinary knowledge and gain an overview of the differing perspectives about a socio-scientific issue. The change made to *DP1* in response to this finding is described in Chapter 7.

Another key finding in relation to *RQ1* was that constructing scientific arguments is likely to be challenging even for those who have experience of engaging in argumentation in other disciplines. The abstract nature of Toulmin's framework proved to be one of the main barriers, with sceptics failing to engage with the activities even when provided with concrete examples of arguments constructed by others because they could not comprehend the abstract terms.

Those in the multiplist and dogmatist clusters spoke of how they had struggled to grasp the warrant, rebuttal, and counterclaim. The reported experiences of the interviewed participants also indicated that they had not grasped the different roles that sources play when constructing arguments, such as providing evidence in support of claims or as backings for warrants. This indicates that the guidance given about how to find information from multiple documents and use this information to construct arguments needs to be better integrated.

Study 1 also revealed how students who lack experience of managing multiple documents often underestimate the time needed to do so. The multiplists had unrealistic expectations about the time needed to synthesise multiple documents; this may have been because they were attempting to use more sources. Poor time-management also proved to be a barrier to critical evaluation, with many of the interviewed participants reporting that they had failed to fully engage in critical evaluation because they had underestimated the time needed. There were also concerns about not having the time to find new sources if critical evaluation revealed that the sources they had chosen thus far were not credible. These findings suggested that intervention should feature an activity that requires students to begin selecting, reading, and critically evaluating sources shortly after they have completed the workshops.

Study 1 enabled *RQ4* to be partly addressed. Students' experiences of and outcomes were explored in relation to the DPs that informed the teaching, learning, assessment, and feedback activities. The findings indicate that *DP5*, the provision of timely positive feedback that motivates students to persist in the effortful processing of belief-inconsistent information, was implemented successfully. All interviewees reported positive outcomes from the feedback they were given during the individual tutorial. This design principle should therefore be retained. There was also evidence that *DP2*, which was to allow students to choose an issue they are interested in, had been also implemented effectively. Five of the interviewed participants demonstrated interest in their chosen issue, though the type of interest differed. Some had personal interests in the topic, others wanted to learn more about the topic to satisfy academic interests or support their career development, and two were interested in how their issue was presented in the media. The

remaining participant offered a more pragmatic explanation; he had chosen his issue because there was plenty of material available. *DP2* should also be retained.

Workshops 2 and 3 need revision if *DP3* and *DP4* are to be implemented successfully. Although all participants arrived at their individual tutorial having already identified at least two of the claims that would feature within their own argument, WS2 was relatively ineffective in teaching students about warrants, rebuttals, or counterclaims: most participants asked for clarification about these components of arguments during their individual tutorial. This workshop therefore needs to be revised to be more effective in teaching students about the components of arguments and how these are combined to create an effective argument. WS3 was not effective in implementing *DP4*, which was concerned with equipping students with the cognitive habits needed to engage with multiple documents. Participants often failed to attend to or engage with the guidance offered about how to assess the credibility of sources and how to critically evaluate content. Participants who struggled to manage multiple documents often experienced negative affect and subsequently adopted maladaptive coping strategies. This workshop therefore needs to provide much clearer guidance about how to critically assess source and content information and make students more aware of the time needed for critical evaluation. It will also be important to teach students about how to skilfully incorporate critical evaluation into their argument.

Despite the mixed success of the first implementation of the design principles, statistical analysis of the differences between the pre- and post-intervention scores indicated there was a tendency for participants to report lower scores for personal justification beliefs after the intervention activities had taken place. Having weaker beliefs in personal justification is thought to be adaptive. It is possible that that this positive change may have been enabled by providing students with a means of finding evidence in support of claims from the sources they were reading. Participants who adopted the strategies suggested in the workshop about reading journal articles effectively and efficiently became less reliant on their personal beliefs to justify their claims.

7 Study 2

7.1 Introduction

Study 2 saw the second iteration of the intervention take place. This study was conducted a full calendar year after Study 1 had occurred. The changes to the design principles and their implementation were informed by the findings from Study 1 (Chapter 6). These findings articulated the challenges that the interviewed participants had faced when trying to construct arguments from multiple documents (*RQ1*) and provided some insights into participants' experiences of and outcomes from the first iteration of the intervention (*RQ4*).

Having implemented the refined intervention, the main purpose of Study 2 was to fully address *RQ4*. As such, the retrospective semi-structured interviews became focused on asking participants to share their experiences of the teaching, learning, assessment, and feedback activities that constituted the intervention. These activities were underpinned by the revised Design Principles. Consequently, Study 2 enabled the effectiveness of these principles to be evaluated. Recommendations for further revision of the Design Principles in future iterations of the intervention are provided at the end of this chapter.

Study 2 was conducted with a new cohort of Science FY students, meaning it was necessary to establish the justification beliefs of these students (*RQ2*) and to assess whether the intervention had successfully promoted the adoption of adaptive justification beliefs.

7.2 Changes Made to the Design Principles and Their Implementation

The design principles had initially been informed by the literature discussed in Chapter 3. Having gathered data in Study 1 which highlighted the challenges faced by the researched students (*RQ1*) and offered a preliminary evaluation of students' experiences of and outcomes from the intervention (*RQ4*), these design principles were revised. The implementation of the design principles was also informed by the findings from Study 1.

7.2.1 Changes to the Design Principles

When reflecting on the findings of Study 1, it became apparent that changes were needed to Design Principles 1, 3, and 4. These changes are described below.

DP1: Provide students with clear guidance about how to select and use appropriate sources when learning new disciplinary knowledge

DP1 had initially been concerned with enabling students to develop disciplinary knowledge and domain specific beliefs. After noting that some of the Study 1 participants had adopted maladaptive search strategies when trying to gain an overview of their chosen socio-scientific issue, this design principle was modified to focus on giving guidance about selecting and using appropriate sources of information when learning new disciplinary knowledge.

DP3: Teach students about the components of arguments and how to find and use information from multiple sources to construct arguments from these components

In Study 1, *DP3* was focused on teaching students how to construct arguments. This design principle was revised to incorporate some aspects that were originally included in *DP4*, which sought to equip students with the cognitive habits needed to engage with multiple documents. The decision to integrate aspects of *DP4* into *DP3* was driven by the recognition that students were capable of reading multiple documents but often struggled to understand what information they needed to be looking for in order to construct arguments.

There was a further change to *DP3*, with rebuttals no longer being conceptualised as being a means of dismissing counterclaims and instead being described as objections to claims. This change was made in response to the difficulties that the Study 1 participants had faced in trying to construct counterclaims and provide rebuttals. It also reflected a shift in my understanding of Toulmin's (1958) argument model.

DP4: Equip students with the cognitive habits needed to critically assess source and content information

The revision to *DP3* for Study 2 necessitated a revision of *DP4*. The focus became narrower, with *DP4* now targeting the cognitive habits needed to critically assess source and content information.

7.2.2 Changes to the Implementation of the Design Principles

The changes to the design principles were implemented through refinement of the workshops and the introduction of an additional assignment, which required students to add annotations to three sources. This assignment aimed to encourage students to adopt the strategies provided in the Controversies in Science and Reading and Evaluating Sources workshops about how to critically evaluate source and content information (*DP4*). It was also hoped that asking students to add annotations to their sources would help students learn how to find information in sources that they could use to construct their arguments (*DP3*). Students submitted their annotated sources ahead of the individual tutorial. The marked assignment was returned at the start of the individual tutorial, giving the tutor the opportunity to supplement their written comments with verbal feedback. Students reflected on this feedback when writing their Reflective Commentary.

All three workshops were refined because of the changes made to the Design Principles, with Workshops 1 and 2 seeing the most extensive changes. WS1, Controversies in Science, was reorganised to implement all aspects relating to *DP4*. This meant the activities designed to teach students how to critically assess sources that had originally been part of the Constructing Arguments workshop (WS2) were now delivered alongside the activities designed to teach about evaluating content information. WS1 also sought to implement *DP1* by introducing students to sources that would enable them to develop their disciplinary knowledge.

WS2 underwent changes to try to enable students to comprehend the components of an argument. The most significant of these changes was that this workshop was split into two sessions. The first of these (WS2a), named Constructing Arguments: Claims, Evidence, and Warrants, aimed to ensure that

students understood the fundamental components of an argument, hence seeking to partly address *DP3*. Toulmin's abstract examples, which had proved largely ineffective, were removed. Changes were also made to the exemplar material featured in this workshop; this was because the scientific ideas associated with the examples used in Study 1 had proved challenging for some students. The new examples, which were about cannabis use and psychotic disorders, acupuncture, Alzheimer's Disease, Chronic Fatigue Syndrome, and consumption of soy protein, were selected because they featured scientific concepts that students would be more familiar with. These examples featured typical and less typical argument structures. The first activity required students to use colour coding to identify the claim and evidence within a typical argument. They were then presented with a less typical argument in which the order of the claim and evidence was reversed. Students were then asked to consider what linked the evidence to the claim in each of these examples, that is, to consider what the warrant would be. The solutions were provided after students had given their suggestions. The second of the Constructing Argument workshops, WS2b, aimed to help students how to better understand how critical evaluating content information, an aspect of *DP4*, would enable them to refine their argument by adding qualifiers and rebuttals to their claims, thus also addressing the second aspect of *DP3*. It also dealt with backings for warrants. This workshop used a subset of the examples that students had met in the previous workshop.

The Reading and Evaluating Sources workshop (WS3), which was concerned with both *DP3* and *DP4* was largely unchanged from Study 1, except for some minor changes to the set of questions relating to critical evaluation. Specifically, the question about why the researchers were interested in the research question was replaced with a question asking about the researchers' authority, expertise, and funding source. In addition, the question about what new knowledge did the study create was removed and the question about whether the findings of the new research matched those of previous research was expanded. Students were also given clearer guidance about the amount of time needed to read and evaluate sources.

7.3 Participants

The Study 2 participants were drawn from a cohort of 34 Science FY students at the researched university. 31 students consented to take part. Of the 30 who completed both the pre- and post-intervention JFK-Q, 20 identified as female and 10 as male. The median age was 19 years, and the mode was 18 years.

7.4 Findings

The main aim of Study 2 was to fully address Research Question 4. This research question, which was concerned with participants' experiences of and outcomes from the intervention, had been partly addressed during Study 1, but there the focus had been on refining the Design Principles and their implementation ahead of the second iteration. Study 2 sought to provide a more comprehensive evaluation. The descriptions of the interviewees not only address *RQ2*, the justification beliefs of the researched students, but also draw attention to whether these participants selected socio-scientific issues that were of interest to them (*DP2*) and sometimes describe participants' prior knowledge (*DP1*) of the issue. There is then further exploration of participants' experiences of and outcomes from the intervention activities, where Design Principles 1, 3, 4, and 5 provide the basis of the thematic framework. The final section reports the results of statistical analysis of the changes to justification beliefs.

7.4.1 Justification Beliefs of the Researched Students

Study 2 participants (*N*=30*)* had relatively weak beliefs in personal justification, moderate beliefs in authority, and relatively strong beliefs in multiple sources. The sample was, nevertheless, heterogeneous, with participants' scores for beliefs in personal justification ranging from 1.33 to 7.33, those for authority ranging from 3.00 to 9.00, and those for multiple sources ranging from 5.40 to 10.00. Table 7-1 displays the pre-intervention justification beliefs of the Study 2 participants.

Table 7-1 *Descriptive Statistics for the Study 2 Cohort*

Sample	*N*	Personal Justification		Authority		Multiple Sources	
		Median	Range	Median	Range	Median	Range
Study 2 cohort	30	4.17	6.00	6.25	6.00	7.60	4.60

Visual inspection of the dendrograms resulting from cluster analysis by Ward's method of the pre-intervention JFK-Q scores suggested two possible solutions: two or four clusters. This dendrogram can be found in Appendix I. Close inspection of the two-cluster solution revealed that the larger of the two clusters was comprised of three small but distinct clusters. Consequently, the four-cluster solution was adopted, with the clusters being named the moderate multiplists, the elevated multiplists, the sceptics, and the dogmatists. The median values for beliefs in personal justification, authority, and multiple sources of participants in these clusters are shown in Table 7-2.

Table 7-2 *Descriptive Statistics for the Study 2 Pre-Intervention Clusters*

Cluster	*n*	Personal Justification		Authority		Multiple Sources	
		Median	Range	Median	Range	Median	Range
Moderate Multiplists	14	4.33	2.33	6.67	2.33	7.80	4.20
Elevated Multiplists	3	7.00	1.00	7.33	2.33	8.40	2.60
Sceptics	5	4.67	2.33	3.67	2.00	7.40	0.80
Dogmatists	8	1.67	1.34	5.17	3.16	6.50	4.20

I had intended to select two interviewees from each cluster, but this was not possible because a relatively low proportion of the Study 2 participants consented to take part in interviews and some of the participants who had consented were subsequently not available. Consequently, there was an uneven distribution of interviewees across the clusters, with three selected from the moderate multiplist cluster, one from the elevated multiplist cluster, one from the sceptics cluster, and

two from the dogmatists cluster. These interviewees had different magnitudes and directions of change in their pre- to post-intervention personal justification beliefs. Where appropriate, the goodness of fit between the cluster name and the participants selected for interview is discussed. The vignettes also describe the socio-scientific issue chosen by each interviewee and explain why this issue was chosen. Where interviewees indicated whether they had prior knowledge of the issue, this information has also been included.

Moderate Multiplists

This cluster was named the moderate multiplists because it grouped together participants with moderate beliefs in personal justification and authority who believed all opinions to be equally valid. It was the largest of the four clusters, with 14 participants. The participants in this cluster held beliefs in personal justification that were comparable to those of the sceptics, but stronger than the dogmatists, and weaker than the elevated multiplists. With respect to beliefs in authority, the moderate multiplists' median values were comparable to the cohort, stronger than the sceptics and the dogmatists, and weaker than the elevated multiplists.

Marc, Mia, and Megan, whose mean scores for their pre- and post-intervention beliefs in personal justification, authority and multiple sources are shown in Table 7-3, were selected as the interviewees from this cluster. Marc was selected because his beliefs in personal justification showed the largest decrease from pre- to post-intervention. Mia also experienced a large change in her beliefs in personal justification, but her direction of change was an increase. Megan had the least change, with her mean score for beliefs in personal justification remaining unchanged from pre- to post-intervention.

Table 7-3 *Pre- and Post- Intervention Mean Justification Beliefs of The Interviewed Study 2 Moderate Multiplists*

Interviewees	Personal Justification		Authority		Multiple Sources	
	Pre	Post	Pre	Post	Pre	Post
Marc	4.67	2.00	6.83	9.00	7.40	7.20
Mia	5.00	7.50	6.83	6.20	7.80	7.80
Megan	4.33	4.33	6.33	7.17	5.80	7.40

All three appeared to view all opinions as being equally valid. Both Marc and Megan initially trusted all sources equally. Furthermore, Marc sometimes appeared to accept statements as fact without questioning whether there was any supporting evidence, though he did recognise that some websites put forward particular viewpoints and tried to persuade their audience to adopt these. Megan and Mia were both keen to represent different perspectives when constructing their arguments. Despite being reminded that she was supposed to be adopting a position in support of one of the perspectives, Megan included points for and against her position because she felt this made her argument more interesting. Mia also incorporated alternative perspectives in her argument but did so through offering rebuttals rather than counterarguments.

Marc was interested in whether powerlifting during adolescence causes growth stunting. He chose this issue because a personal trainer had told him that powerlifting exercises can impact a child's growth. Marc regularly exercised at the gym and was hoping to progress to an undergraduate degree in sport rehabilitation, so this issue was of personal interest and related to his career goals. Although he had not read beyond sports publications aimed at the public, he was aware that there were differing perspectives on the issue. He initially believed that it was likely that powerlifting could cause growth stunting but changed his mind when he was unable to find any scientific evidence in support of this position.

An aspiring Physiotherapist, Mia wanted to explore an issue that related to her future career. She decided to choose an issue that she knew little about because this would allow her to approach it with an open mind and reach a decision

based on what she had learnt. She initially planned to compare the effectiveness of myofascial release and static stretching on performance by writing a balanced argument that presented the evidence for both. After being reminded that she needed to take a position, she opted to argue in favour of myofascial release as the experimental results indicated this form of treatment was more effective than static stretching.

Megan opted to explore whether banning supermarkets from having sugary snacks near the check-outs would help tackle obesity. She chose this issue because she was interested in nutrition and intended to pursue a career in this field. Megan was interested in the different viewpoints she encountered when reading research about dietary intake and obesity, though struggled to find studies that were directly related to the issue she had selected as the focus of her argument.

Elevated Multiplists

The participants in this cluster had pre-intervention scores of similar magnitudes for all three justification beliefs, suggesting that they were multiplists. Their scores for beliefs in personal justification and authority were, however, much higher than those of the moderate multiplists, hence they were sorted into a distinct cluster and termed the elevated multiplists. Elliot, whose mean score for beliefs in personal justification showed a large decrease from pre- to post-intervention, was selected as the interviewee from this cluster. Table 7-4 displays his pre- and post-intervention scores.

Elliot's elevated multiplist beliefs appear to have been heavily influenced by what he had read and the enjoyment he experienced when considering alternative perspectives. For instance, he referred to the works of Yuval Noah Harari and indicated that he regularly read MIT Technology Review, The Economist, Financial Times, and Nature. He also spoke of how he enjoyed discussing controversial issues and how he had gained experience of doing so when studying Psychology.

Table 7-4 *Pre- and Post- Intervention Mean Justification Beliefs of the Interviewed Study 2 Elevated Multiplist*

Interviewee	Personal Justification		Authority		Multiple Sources	
	Pre	Post	Pre	Post	Pre	Post
Elliot	7.33	3.67	6.67	7.83	7.40	6.20

Elliot chose to explore the issue of whether there should be a moratorium on gene editing. He was driven by his fascination with the possibilities offered by germline gene editing; hence his interest was primarily academic. Even though most researchers are currently opposed to germline gene editing, and he was aware of the risks of misuse, Elliot did not allow himself to be concerned at the possible implications of germline gene editing because he believed its use to be inevitable. Elliot had strategies in place for reading sources and appeared to already possess the cognitive habits needed to critically evaluate source and content information.

Sceptics

When Greene et al. (2008) used the term sceptics, it referred to those with strong beliefs in personal justification. Although better described as moderate rather than strong, particularly as they were comparable to the overall cohort value and those of the moderate multiplists, the personal justification beliefs of the sceptics were higher than those of the dogmatists. According to Greene et al., sceptics should also differ from dogmatists with respect to beliefs in authority. The median value of authority for participants in this cluster was lower than the dogmatists. In fact, it was lower than all other clusters and the cohort.

Seb was selected from this cluster. As shown in Table 7-5, his mean score for beliefs in personal justification did not change from pre- to post-intervention.

Table 7-5 *Pre- and Post- Intervention Mean Justification Beliefs of the Interviewed Study 2 Sceptic*

Interviewee	Personal Justification		Authority		Multiple Sources	
	Pre	Post	Pre	Post	Pre	Post
Seb	4.33	4.33	3.83	5.33	7.60	6.40

Seb was in the sceptic cluster because he had a relatively weak mean score for pre-intervention beliefs in authority. These may have resulted from him being unsure about who to trust. Despite describing the NHS as being unbiased and seeking to provide facts, he commented on how NHS advice sometimes differed from the recommendations made in recent scientific research. It was also evident that he feared that online organisations were trying to sell something, leading him to distrust what he read on their websites.

Seb decided to explore whether hormone replacement therapy (HRT) causes venous thromboembolism after talking to his family member about issues that were controversial within the field of general practitioner (GP) medicine. He knew little about the issue beforehand, other than this was an issue where GPs need to be able to assess both the benefits and risks for their patients in order to provide appropriate advice. Seb expected there to be consensus about whether the benefits outweighed the risks but instead found differences of opinion. He ultimately took the position that the risks did not outweigh the benefits for women with relatively mild symptoms. Having previously gained experience of using multiple sources during a previous course, Seb felt that he had effective strategies in place for using sources to construct an argument.

Dogmatists

The dogmatists' beliefs in personal justification were much weaker than those of the participants in the other clusters. Darcey and Danielle were selected as the interviewees from this cluster. Their pre- and post-intervention scores are shown in Table 7-6. Darcey's mean score for beliefs in personal justification changed the least, showing a small decrease, though she was already close to the minimum

possible value at pre-intervention. Like Darcey, Danielle's pre-intervention mean score for beliefs in personal justification was just 1.33. Danielle was chosen because she had the largest increase in personal justification beliefs from pre- to post-intervention.

Table 7-6 *Pre- and Post- Intervention Mean Justification Beliefs of the Interviewed Study 2 Dogmatists*

Participant	Personal Justification		Authority		Multiple Sources	
	Pre	Post	Pre	Post	Pre	Post
Darcey	1.33	1.00	5.17	5.50	9.40	9.00
Danielle	1.33	3.00	4.83	4.67	9.60	9.80

Darcey selected an issue that she was personally interested in when she chose to argue that animal testing should only be used in situations where is no other alternative. She was initially reluctant to acknowledge that others may hold views that differed to hers. Moreover, she perceived that her position was aligned with that of many researchers, saying that most were against animal testing, though she did also say that she had found a range of perspectives in the literature. Darcey was aware that she was dealing with an emotive issue and felt that she was not alone in allowing her emotional response to influence her viewpoint. She later reflected that she felt she had become more objective in how she assessed the evidence.

Danielle chose the issue of whether elective Caesarean sections were problematic. Describing herself as having strong beliefs in women's rights and bodily autonomy, this issue challenged Danielle to reconcile her pre-existing personal values with the findings from scientific research. Finding herself in agreement with scientific research about the risks of elective Caesarean sections, Danielle realised that approaching the issue from a different angle would enable her to argue against elective Caesarean sections without compromising her deeply held values. Danielle described how she had prior experience of writing essays when studying English and was familiar with some of the components of an argument, though also noted that the structure of a scientific argument differed considerably to that of an English essay.

7.4.2 Participants' Experiences of and Outcomes from the Intervention Activities

The reported experiences of the interviewed participants and their outcomes from the intervention activities were analysed in relation to the design principles.

DP1: Provide students with clear guidance about how to select and use appropriate sources when learning new disciplinary knowledge

In Study 2, this design principle informed teaching and learning activities that took place during WS1. It is difficult to assess whether the workshop was effective because only one of the interviewees commented on their experience of being given guidance about different sources. Darcey explained she had found it useful to learn about meta-analyses during the Controversies in Science workshop because she had used a few of these when constructing her argument.

The individual tutorial proved to be a more effective means of providing students with guidance about the types of sources they should be using when constructing their argument. Almost all interviewees discussed types of sources and their suitability during their tutorial. These discussions proved effective, with post-tutorial reflections indicating that participants had understood the guidance they had been given and were planning to implement the advice. For example, Mia talked of using the Introduction sections of journal articles as a source of information for her Introduction and Darcey described how gathering sources that provided an overview of the claims she intended to make within her argument had allowed her to "immediately establish clear arguments".

The individual tutorial may have proved effective because students had submitted annotated sources in advance of the tutorial. This allowed the tutor to understand what sources students had been accessing and how they were intending to use these sources in their assignment. All the interviewed multiplists had selected at least one secondary source, such as a review article, systematic review, meta-analysis, or commentary. Consequently, the guidance given during the individual tutorials tended to be about the difference between primary and secondary sources, and how such sources should be used when constructing an argument. For example, Mia had selected a systematic review and had initially

planned to use this as evidence. During her tutorial, Mia was given guidance about the difference between research articles and review articles, systematic reviews, and meta-analyses. Similarly, Megan had read a commentary from a fact-checking website that assesses whether media reporting of recent research accurately represents the research but had not understood that the article's authors had not carried out the research. When reflecting on the feedback she was given in the tutorial, Megan said that she was now aware she needed to "pick primary resources (*sic*) instead of secondary", suggesting that the discussion had proved effective in enabling her to understand the difference. Marc had also selected a source that was readily available to the public rather than one written for the scientific community. He explained that he had selected general health websites because they "had the most relevancy in terms of addressing my issue".

Not all interviewees required guidance about choosing suitable sources. Elliot, Seb, and Danielle were all regular readers of scientific publications, referring to Nature, The British Medical Journal (BMJ) and The Lancet respectively. These participants were largely comfortable in using primary sources to develop their knowledge and understanding, though Danielle did acknowledge that her strategy of using primary sources to gain an overview was not effective at first because her search was too broad:

> "I looked at relevant themes, the introductions, and conclusions researchers had drawn. (This was) useful for gaining a general overview for my introduction but at that point I had no real focus so (it was) mildly ineffective".

In summary, the discussions within the individual tutorials about sources proved effective in enabling those who lacked prior experience of reading scientific publications to select and use suitable sources of information to develop their disciplinary knowledge and construct an argument. The lack of comments relating to the Controversies in Science workshop makes it difficult to draw conclusions about its effectiveness in implementing *DP1*, though it does appear that the guidance about how to assess the credibility of websites was not attended to by all.

DP2: Allow students to choose topics they are interested in

This design principle was primarily implemented during WS2, when students were shown examples of the socio-scientific issues that had interested previous Science FY students. Students were subsequently asked about their choice of issue during their individual tutorial. Although none of the interviewees commented on whether seeing examples of previous students' work had influenced their own choice of issue, all had selected issues which they were interested in, indicating that this design principle had been implemented effectively.

DP3: Teach students about the components of arguments and how to find and use information from multiple sources to construct arguments from these components

In Study 2, WS2a and WS2b sought to teach students about the components of arguments. These workshops featured multiple examples of arguments in various stages of completion. The example arguments also differed in how the components were organised within the text, with some having the claim first followed by the warrant and then the evidence and others having a less typical structure. Students were encouraged to use colour coding to highlight the different components, as this would enable them to compare how arguments were structured.

Almost all the interviewees spoke about their experiences of and outcomes from these workshops. For instance, Megan explained that she found the use of colour coding to highlight different components to be helpful when constructing her own argument:

> "When each individual claim warrant was in the different colour that was helpful because then you could apply that to your own work and when I was doing mine I was highlighting the claim and everything else and different colours to ensure that I've put them in each paragraph".

Seb also referred to the use of colour coding as being a positive experience, saying "the colour coding definitely helps".

When asked about the multiple examples, most interviewees reported finding these to be useful. For example, Danielle commented, "I think the examples that we looked at for everything were definitely the most helpful part". Megan explained that she had returned to these examples when constructing her own argument:

> "I'd go back to the PowerPoint and see how they'd done it and then go back to my own work and think well I have added that and you know yeah it was really helpful to have the guidance".

There was, however, one interviewee who was not as keen on having multiple examples. Seb remarked this was confusing for him, though he went on to say he had looked back at the examples when writing his assignment. This suggests that his initial response had not subsequently prevented him from engaging with the resources.

Many of the interviewees spoke of returning to the workshops when constructing their own arguments. Although this indicates that they found the guidance to be helpful, it also suggests that they had not immediately grasped how to construct and refine arguments. Indeed, many needed further guidance during their individual tutorials. Of these, Marc seemed to be the least aware of the components of an argument, and consequently was given guidance about much of what had been covered in the workshops. Although he said he found it "beneficial" to have been given what he variously described during his tutorial as a "formula", "foundation point" and "groundwork" for constructing arguments, he later said "I think I got a bit confused", indicating that he had not fully grasped how to construct an argument. Another of the moderate multiplists, Megan, also needed clarification of the claims, evidence, and warrant during her tutorial, though this was primarily because she was unsure about how to use information from sources to construct her argument:

> "Are you describing what they've done … all their facts and stuff like that and then do you say well because of this I believe or are you literally being them in a sense".

Megan, Mia, and Darcey were all given guidance about qualifiers and rebuttals. These discussions had different outcomes, with Darcey remaining confused about the difference between these components. Megan may also have misinterpreted some of the guidance, though this may have been because she was keen to include the opposing view. Mia had a much more positive outcome, quickly realising that selecting flawed studies as rebuttals would allow her to dismiss any objections to her claims, asking, "so could you … be like ooh as a rebuttal this person said this but then you could also analyse the source that has come from and then try and almost disprove the rebuttal".

Three of the interviewees, Elliot, Seb, and Danielle, spoke of having gained experience of constructing arguments during their pre-university education. Of these, Elliot and Danielle did not ask for clarification about any of the components of an argument during their tutorials, though Danielle later spoke of how it had taken her a while to get to grips with the backing. Seb asked for help with the qualifier, though discussions during his individual tutorial indicated that he had identified limitations of his claim. It therefore appears he had grasped the concept of a qualifier without knowing the associated term.

To summarise, most interviewees reported positive outcomes from using the resources associated with the Constructing Arguments and Refining Arguments workshops, indicating that the resources had aided them when constructing their own arguments. The learning activities during these workshops did not, however, fully succeed in teaching students about the components of arguments, with many needing to seek clarification during their individual tutorials. For some, these discussions brought clarity whereas others remained confused.

Having been revised to incorporate some aspects of what had previously been in *DP4*, *DP3* also informed the design of the workshop about Reading and Evaluating Sources (WS3). Five interviewees commented on the guidance they had been given in this workshop. Three of these were those with prior experience of constructing arguments: Elliot, Seb, and Danielle. These participants differed in their responses. Elliot described himself as being "not very keen at all" on the activities featured in this workshop. Seb initially took a neutral stance in his response, saying it had reinforced what he already knew, though he also remarked

that he could "understand if people have not done it before its very very helpful". Danielle's response was much more positive, describing how the strategies for reading sources could help her spend less time reading sources:

> "That exercise where we went through a journal article and said right these are the bits that are going to be the most helpful and this is going to help you just spend less time on reading that was really helpful to me because when you've got a sea of words in front of you and you're just going I have 15 journal articles to read they're all 100 pages long please somebody help me it's a very useful technique to have definitely".

It does appear, however, that Danielle did not fully implement the strategies suggested during the workshop when first reading sources; within her Reflective Commentary, she wrote that she had read the whole paper to get more information. Danielle's approach to reading changed over time, and she had a very positive experience of looking for information to use within her argument:

> "I absolutely loved having to say that's the data I'm using because it just stopped me from going down that whole spiral route of oh that's interesting so I feel that it helped keep me focused which was really good".

Mia appears to have had a similar experience to Danielle because she also initially read the whole article "just to make sure that I wasn't missing out on things". Yet she also recognised that she needed to be reading with a purpose in mind. She later commented on how the workshop activities had helped her stay focused when looking for information to use within her argument:

> "Having the workshops (you) just know in your head what you don't need and what you do … it just gives you more of a focus on what you should be looking at and I think that definitely helps".

Darcey differed from Mia and Danielle in that she appears to have immediately adopted a very focused approach. She reported reading to find the main points, gain understanding of key information, and determine the relevance of the source for her argument. It does, however, appear that Darcey may have misunderstood that she would need to critically analyse her sources to form qualifiers and rebuttals for her arguments and instead seemed to think that these components would be found within the sources. Nevertheless, Darcey's experiences, along with those of Danielle and Mia, indicate that at least some of interviewed participants had been receptive to and acted upon the reading strategies suggested in the workshop. There was also evidence that some participants had paid attention to the idea of using a template or checklist when reading sources to help guide their reading and focus on key information. For example, Seb reported he had used a checklist when reading sources. Likewise, Danielle indicated that she was planning on using a checklist to aid her in selecting relevant information and later described how she had done so when reading sources to help her monitor her progress. Although it was unclear whether Seb and Danielle had used the template provided in the workshop or had created their own, these findings indicate that the guidance given in the workshop enabled participants to develop effective strategies for finding information in sources.

Overall, the Reading and Evaluating workshop enabled participants develop their understanding of how to use information from sources to construct their arguments. The two interviewed dogmatists, along with one of the moderate multiplists and the sceptic, all reported using strategies and resources from this workshop. It seems that this workshop, like those on Constructing Arguments and Refining Arguments, proved most useful once participants were engaged in constructing their own argument.

The assignment activity which involved students annotating sources was met with a mixed response. Seb and Elliot, who both had previously studied science subjects, did not see the value in writing annotations and preferred instead to use their own note-making strategies. Their honest reflections were welcomed, not least because it later became apparent that all the interviewed participants had adopted alternative strategies for making notes when reading sources. Like Elliot,

Seb had reverted to a strategy he had used previously. Mia moved away from highlighting text on the source document and instead created notes to record the key points. She subsequently used these notes to match claims to evidence, warrants, and rebuttals. Even Danielle, who described adding annotations about how the information would be used in an argument as "just a great way of streamlining the whole process", explained that she had decided to make notes rather than continuing to annotate sources. Both Mia and Danielle made notes because they did not want to have to search back through their sources when trying to find information to include in their arguments. As Mia rightly noted:

> "If you're writing your claim and then you go oh but what evidence was it you've got to flick through 70 odd pages or however much it was to try to find it again but if you've got it there it just makes life a lot easier".

This was an astute observation and one that caused me to further reflect on how best to assess students' competence in identifying useful information from sources without requiring them complete a task that may hinder their ability to construct a coherent argument.

Despite its limitations, the annotation activity was largely effective in achieving its aim of encouraging participants to think about how they would use information from sources when constructing their argument. For instance, Seb described how the annotation activity had given him a "starting point and a "means of getting baseline information" which he subsequently used in his assignment. Megan also thought of her annotations as a starting point:

> "It gave almost a start to the writing an argument assignment it helped immensely with that because you had a basic structure for it you had all the knowledge that you needed to know and it was just a case of now structuring it into an argument".

Darcey reflected on how she would have made notes on her sources even if she had not needed to do so for the annotation activity. She thought this was "the most

logical thing to do" because it allowed her to record "this is what it's saying in relation to the argument".

As almost all interviewees reported positive experiences of the annotation activity, it appears that it was successful to some extent, not least because it provided the starting point for conversations in the individual tutorial. It was, however, also apparent that participants quickly adopted different ways of making notes from their sources.

DP4: Equip students with the cognitive habits needed to critically assess source and content information

DP4 was partly implemented in Controversies in Science workshop (WS1), when students were asked to compare the credibility of websites by answering a set of questions which asked about the currency, reliability, authority, and purpose of the website. The Reading and Evaluating Sources workshop (WS3) saw these questions, alongside others, being included in a template that students were advised to complete when reading a source. This workshop also featured an example of an annotated source.

The activities within WS1 and WS3 and the resources associated with them were well-received by some of the participants who lacked any prior experience of assessing source information. For instance, Mia and Darcey both spoke about how they had found the website credibility activity helpful because it had given them criteria against which to judge other websites. For example, Mia said: "I think it helped a lot … because it made you more aware of what the good reputable websites would say … what you were looking for".

While some had clearly attended to the guidance they were given about how to critically assess source information, others had not. Marc appeared to have been unaware of the need to critically evaluate source information. After discussing his choice of sources in his tutorial, he reflected on how he now understood that his source selection was not appropriate because it had "lead (*sic)* me to choosing sources that weren't from scientific organisations or from people with scientific backgrounds". He further commented that he was going to look instead for "articles written by professionals rather than chatrooms, for example,

where anyone can express their opinion". It therefore appears that Marc had initially been unaware that not all sources are equally credible.

Participants who had prior experience of critically evaluating sources differed in their responses to the workshops and their associated resources. Danielle, who was changing disciplines having previously studied the Arts, embraced the opportunity to learn how to critically analyse scientific sources. She described how her evaluation checklist, which reminded her to check the publisher, authors, funding, and conflicts of interest, had been "an excellent way to focus". Danielle was also very appreciative of being given an example of how to annotate sources, saying it helped her understand how to assess the credibility of an author. In contrast, Elliot admitted he was relying on using the prestige of the publisher as a proxy for credibility. Despite knowing that this was a flawed approach, he later admitted that he had persisted with his "good journal good authors" approach to assessing source credibility.

Even when participants had engaged with activities and reported finding them helpful, it appears that the workshops were not entirely effective in embedding sourcing behaviours. Five participants were given additional guidance during their individual tutorial. Marc was given advice about how to assess the quality and credibility of websites. Seb was unsure about what constituted recent research. Megan, Darcey, and Danielle, who had all selected secondary sources such as reviews and commentaries, were unsure of how to critically evaluate such sources. The two participants who did not require guidance during their tutorials already had well-developed sourcing behaviours. Elliot was able to articulate how the lead authors of the sources he had selected had contributed to research. Mia had been able to decide which perspective to adopt in her argument after noticing that there were two distinct groups of researchers. She described the group she sided with, who were sports health professionals, as being the 'specialists' and the other, who were physiotherapists, as being 'generalists'.

For almost all participants, the guidance provided during tutorials resulted in positive outcomes. Marc switched to using peer-reviewed, primary sources, and routinely checked the date of publication and the author's qualifications. He later reflected "I guess that was one of the things in this assignment wasn't it, it was

learning who to trust". Having initially believed all sources to be equally valid, Marc now recognised that credible sources are written by those with expertise or who are affiliated with a professional organisation. This shift in understanding was mirrored by the change to his justification beliefs, which saw him placing less trust in personal justification and more trust in authority. Megan also described how learning to pay attention to an author's expertise had fundamentally changed her view of what constitutes evidence:

> "A lot of people in mine in particular worked in hospitals and worked in unis and I think that was quite the turning point for me it wasn't just a general viewpoint from a parent or a person off the street it was someone with qualifications and a higher level".

While Megan's scores for beliefs in personal justification did not change, she did have stronger post-intervention beliefs in justification by authority and multiple sources.

After receiving positive feedback about the amount of effort she had been putting into checking author's credibility, Darcey continued to invest a lot of time and effort into assessing source information. In stark contrast, Seb scaled back his efforts after getting positive feedback about engaging in critical evaluation. Seb had been checking funding sources, conflicts of interest, qualifications, and affiliations of all authors. After his tutorial, he decided that he would limit his efforts to the lead researcher to save time, saying, "When critically evaluating spend less time reading through all the researches (*sic*) and just stick to the main researcher and check their credibility". Although somewhat unclear, it is possible that he used a passing comment about the lead author being the one who secures the funding as justification for his decision to allocate less time to check the authors' credentials.

In summary, these findings indicate that participants' responses to the intervention were more dependent on their prior experiences of assessing sources than their justification beliefs. Participants who were already competent in evaluating source information sometimes opted out of some aspects of the intervention but were not disadvantaged by doing so. Those who lacked experience in evaluating scientific sources reflected on how the guidance about how to

critically assess the credibility of a source had brought about changes to their understanding of what constitutes expertise and evidence in science.

DP4 was also concerned with equipping students with the cognitive habits needed to critically evaluate the content of sources. This design principle was addressed during WS3, when students were asked to critically evaluate a piece of scientific research by answering a series of questions within a template. Seb, Darcey, and Danielle all reported finding the template provided during WS3 to be very effective when assessing their own sources. Danielle was particularly enthused about learning how to critically evaluate scientific research:

> "I came from an English background when English people say analyse and when Science people say analyse they mean very different things … it was a really good process to go through … I feel it was a really good learning experience".

Seb reported that he was paying attention to "how relevant the studies were and who they were representative of", though he did express concern that doing so was taking him a significant amount of time. In contrast to the somewhat superficial approach to critical evaluation adopted by Seb, Darcey was very thorough in how she assessed the research, paying attention to "sample sizes, appropriateness of techniques, repeatability, experimenter bias" as well whether extraneous factors had affected the results. She recorded all this information into a table and used it to decide whether to use the source in her argument and where to use it. Although none of the moderate multiplists reported using the template, it appears that they benefitted from being taught how to analyse content information. Marc and Megan began with a very limited understanding of critical evaluation. During their reflective interviews, it was evident they had developed some understanding, though they still had more to learn. For instance, Marc spoke of how he now recognised that the sample size of 12 may not be representative. Megan also focused on the sample as a means of assessing the quality of the evidence, though needed guidance about what is meant by the target population as she assumed that all research needs to be generalisable to the whole population. Mia used multiple measures to critically assess scientific research. She described how she paid

attention to sample sizes, who the participants were, and whether the data collection method was valid:

> "The sample size and the type of things they were doing … whether it was for long enough so people would be stretching but it depends on the amount or type of exercise that they've done beforehand and how long they were stretching for … if it wasn't that great then it led you to believe that it hasn't been carried out in a very effective way … if they got a bigger sample size then it would probably have been a different result … if you've got two people who are very healthy or athletic and they both happen to do better after stretching then that argument is gonna be oh well stretching is better but if you had sixty and it was more like two preferred stretching than the rest then you're going to get a different result".

One of the outcomes of the intervention for Mia was that she came to realise that she could use critical evaluation as a means of dismissing rebuttals to claims if she selected studies that were easily criticised. In doing so, she was able to include what she described as being 'bad' studies as well as those which were 'good'. Mia felt this was important as she did not want to create the impression that "every experiment is a good experiment".

Unlike all other interviewees, Elliot appears to have paid little attention to the studies' participants or the data collection methods used and instead evaluated the research he used within his argument on the basis of which perspective the researchers held. It therefore appears that for Elliot, this aspect of the intervention was not effective in achieving its intended outcomes.

These findings indicate that participants who lacked experience of critically evaluating scientific research became more competent in doing so, though their levels of competence differed. This may be partly due to the extent to which they adopted the suggested criteria. The dogmatists appear to have paid the closest attention to the guidance given, and as a result they used the most criteria when assessing studies. They were followed closely by one of the moderate multiplists,

Mia. While Seb, a sceptic, reported finding the checklist helpful, it appears that he had underestimated the time needed to engage with all aspects of critical evaluation and so chose to focus on just two issues. This placed him close to the other moderate multiplists, Marc and Megan, in the extent to which he enacted the advice given during the workshop. Elliot largely ignored the advice provided in the workshops and did not critically evaluate the research, focusing instead on the authors' position in the debate.

DP5: Provide positive, timely feedback which motivates students to engage in effortful processing of belief-inconsistent information

Design Principle 5 was unchanged following Study 1, where the individual tutorials had proved highly successful in providing positive, timely feedback which motivated students to engage in the effortful processing of belief-inconsistent information. The only change made to the implementation of this design principle for Study 2 was to require students to submit annotated sources prior to their tutorial. As a result, the individual tutorial often began with the tutor giving feedback about their choice of sources and the annotations on them. Students were still able to ask for guidance about how to use sources to construct arguments and other related matters.

When reflecting on their feedback, six of the seven interviewees reported that receiving feedback had been a wholly positive experience. Marc and Darcey both described their feedback as being constructive, with Darcey explaining that this was because it had made her "re-evaluate what sources I might use as qualifiers / rebuttals (*sic*)". Seb similarly related the feedback to particular components; in his case it was qualifiers and backings. Megan wrote in more general terms about how she had found the feedback helpful in constructing her argument: "The improvements helped my techniques and practices which will help me write a successful argument".

Megan, Mia, and Danielle all expressed positive affective responses to the feedback, with Megan saying that it motivated her and Mia describing how she felt supported. She further explained that this was because the feedback had reassured her: "The feedback made me feel supported in that I was on the right track in my

analysis and helped me focus my ideas on how I should write the argument". Danielle was also reassured by the feedback she received. She wrote of how she was happy to hear that she had conveyed her ideas clearly, both in relation to how she was planning on using sources and the sense she had made of the issue. She was also pleased to know that she was critically evaluating correctly. In addition, she expressed gratitude for the advice she was given about selecting suitable sources, saying that she was "grateful for this learning experience".

As indicated earlier, Elliot was not afraid of being critical. While he agreed that his annotations lacked detail, he did not agree with the advice given, writing: "I disagree that deeper annotations would allow a clearer quality of thinking as I already had summaries produced elsewhere". Elliot's ambivalent response set him apart from the other interviewed participants, who instead described the feedback as being "useful", "helpful", and "constructive". Many also described how they had experienced positive affect. Outcomes from the feedback were also largely positive, with participants describing how they had developed better understanding of what constitutes a suitable source, overcome misunderstandings about components of arguments, and learnt where to find information relating to these in sources.

7.4.3 Changes to Participants' Justification Beliefs

The difference between the pre- and post-intervention scores was used as a quantitative outcome measure. Table 7-7 shows the descriptive statistics for the pre- and post-intervention justification beliefs.

Table 7-7 *Descriptive Statistics for Participants Pre- and Post-Intervention Beliefs in Study 2*

Justification Beliefs	Pre-Intervention				Post-Intervention			
	Mean	SD	Median	Range	Mean	SD	Median	Range
Personal Justification	3.88	1.67	4.17	6.00	3.80	1.92	3.33	7.00
Authority	5.90	1.44	6.25	6.00	6.27	1.49	6.19	6.33
Multiple Sources	7.68	1.19	7.60	4.60	8.11	1.22	8.10	4.60

At the cohort level (N = 30), the mean and median values for beliefs in personal justification decreased from pre- to post-intervention. In contrast, mean and median values for beliefs in justification by multiple sources increased.

The differences between pre- and post-intervention scores were subjected to statistical analysis using one-tailed Wilcoxon matched-pairs signed-ranks test with the Pratt's correction for tied values. Three paired t-tests were also conducted. Table 7-8 displays the results for the comparisons between pre- and post-intervention scores for the justification beliefs.

Table 7-8 *Statistical Analysis of Differences Between the Study 2 Pre- and Post-Intervention Scores by Justification Belief*

Justification Beliefs	*W*	*p*	*t*	*p*
Personal Justification	18.00	0.4295	0.253	0.401
Authority	156.0	0.0552	1.68	0.052
Multiple Sources	164.0	0.0459*	2.20	0.018

These comparisons indicate that the pre- and post-intervention scores for personal justification beliefs were not significantly different. The change to beliefs in justification by multiple sources from pre- to post-intervention was statistically significant at the p=0.05 level. It can therefore be concluded that participants' beliefs in justification by multiple sources became significantly stronger. With regards to beliefs in authority, although both statistical tests suggested that the change was approaching significance, the direction of change was unclear because the mean value increased whereas the median value decreased. It has therefore been concluded that the changes to pre- and post-intervention scores for beliefs in authority were not interpretable.

The changes to participants' scores were further analysed by summing the justification beliefs across the three sub-scales, having first rescaled the values for personal justification. The pre-intervention summed value was compared to the post-intervention summed value using the Wilcoxon matched-pairs signed-ranks test for differences. The Pratt's correction was not needed as there were no tied values. It was, however, necessary to adjust the significance level (α) using the Bonferroni method to allow for multiple comparisons using the same data sets. As a result, the p value could only be judged to be statistically significant if equal to or below 0.025. Despite this more conservative measure, there was a statistically significant increase in participant's summed justification belief scores from pre- to post-intervention, W (30) = 192.0, p = 0.0239. This outcome indicates that participants' justification beliefs became more adaptive.

Participants' post-intervention justification beliefs were also used to sort participants into clusters using the same approach as for their pre-intervention scores. The dendrogram resulting from Ward's method of cluster analysis was subjected to visual inspection and it was determined that there were two clusters; the descriptive statistics for these clusters are described in Table 7-9. **Table 7-9**The largest cluster (n=23) was named the dogmatists because these participants had relatively strong beliefs in authority but relatively weak beliefs in personal justification. This cluster included all those who had been in the pre-intervention

dogmatists cluster (*n*=8). It also included 11 of the 14 moderate multiplists, two of the three elevated multiplists, and two of the five sceptics.

Table 7-9 *Descriptive Statistics for the Study 2 Post-Intervention Clusters*

Justification Beliefs	Dogmatists (n=23)		Multiplists (n=7)	
	Median	Range	Median	Range
Personal Justification	3.00	3.33	7.00	3.00
Authority	6.17	4.67	6.20	5.83
Multiple Sources	8.20	4.60	7.80	2.80

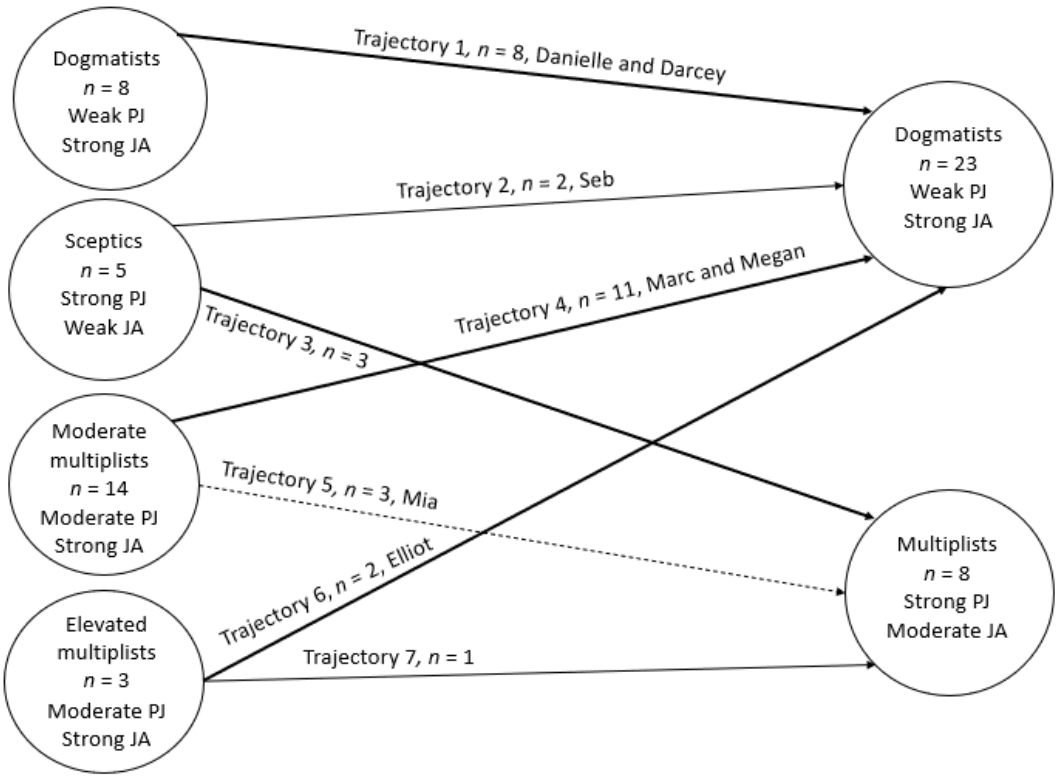
Of the 11 moderate multiplists who became dogmatists, nine reported lower personal justification beliefs post-intervention; the other two had no change in their scores. The two elevated multiplists who became dogmatists had lower post-intervention scores for personal justification, as did one of the sceptics; the other sceptic had no change. Interestingly, the personal justification scores for six of the eight dogmatists increased, leaving just two dogmatists whose reported personal justification beliefs decreased.

The smaller cluster (*n*=7), named the multiplists, consisted of three of the moderate multiplists, 1 of the elevated multiplists and three sceptics. All these participants reported higher scores for personal justification beliefs post-intervention.

The movement of participants from pre-intervention to post-intervention clusters led to there being seven different trajectories, which are illustrated in Figure 7-1. The interviewed participants can be found on five different trajectories. Danielle and Darcey both followed Trajectory 1, remaining as dogmatists. Seb, Marc, Megan, and Elliot all became dogmatists, following trajectories 2, 4 and 6 respectively. Mia followed trajectory 5, indicating that she had remained a multiplist, albeit one with stronger beliefs in personal justification but more moderate beliefs in justification by authority.

Figure 7-1

Trajectories from Pre- to Post-Intervention Study 2 Clusters



Note. Changes in cluster membership before and after the intervention, with number of participants moving from each pre-intervention cluster to each post-intervention cluster. Dashed lines indicate that the proportion of students moving along that the trajectory was less than 25% of those in the pre-intervention cluster, thin solid lines indicate the proportion is between 25% and 50%, and thick solid lines indicate the proportion is at least 50%. The names indicate the interviewees who followed that trajectory.

7.5 Summary

Study 2 participants differed in their pre-intervention justification beliefs (*RQ2*). The largest cluster, the moderate multiplists, and the smallest, the elevated multiplists, had median values for personal justification and authority which indicated they believed all opinions to be equally valid. These clusters differed,

however, in the magnitude of their scores for beliefs in personal justification, with the elevated multiplists reporting stronger beliefs. The sceptics differed from the dogmatists with respect to both beliefs in personal justification and authority.

Analysis of the differences between the pre- and post-intervention summed measure of justification beliefs indicated there had been a statistically significant increase, suggesting that participants' overall justification beliefs had become more adaptive. One of the component measures, beliefs in multiple sources, had also increased significantly.

Study 2 saw three of the five design principles being revised in response to the findings from Study 1. *DP1* became focused on the providing students with guidance about appropriate sources to use when learning background knowledge. *DP3* brought together learning about the components of arguments with learning how to find these components when reading sources and *DP4* became focused on analysing source and content information. *DP2*, which allowed students to choose topics they are interested in, and *DP5*, which encouraged the provision of positive feedback, remained unchanged.

Analysis of interviewed participants' experiences of and outcomes from the implementation of these DPs through the intervention activities enabled *RQ4* to be addressed. DP2 proved effective: all participants selected socio-scientific issues they were interested in, though it was unclear whether seeing examples of issues chosen by previous students had influenced topic choice. DP5 also continued to be effective, with feedback almost always leading to positive affect.

When implementing the revised DP1, the Controversies in Science workshop tried to place greater emphasis on analysis of source information. It is difficult to assess the effectiveness of this change because the interviewees rarely commented on their experiences of this workshop. What is apparent is that the individual tutorial played an important role in providing an opportunity to give advice about what constitutes a suitable source.

Interviewees' reported experiences indicate that neither the Constructing Arguments and Refining Arguments workshops nor the individual tutorials were entirely successful in enabling students to get a firm grasp on the components of arguments, an aspect of *DP3*. Interviewees reported returning to the examples

provided when constructing their own arguments. They also often mentioned how they had adopted the colour coding used within these examples to highlight the different components of their own arguments.

DP3 was also concerned with teaching participants how to use sources to construct arguments. The annotation activity appears to have been successful in providing a starting point and helped the tutor identify those who were having difficulties understanding what constitutes a suitable source or what the different components of an argument are. Yet none of the participants opted to continue adding annotations to their sources, indicating that this approach had limited utility.

The revised version of *DP4* saw the focus shift to equipping students with the cognitive habits needed to critically evaluate source and content information. The sceptic and the elevated multiplist, both of whom had prior experience of working with multiple documents, did not engage with the guidance about critically evaluating source information. Those who did, including a dogmatist with prior experience, had positive experiences of doing so. Some of those who lacked prior experience even described their experiences as being transformational.

The effectiveness of the guidance about critically evaluating content information varied according to participants' justification beliefs, with dogmatists paying attention to the most factors when critically analysing research. The dogmatists had also adopted the criteria for critically evaluating sources, suggesting that they were the most receptive to being given advice about how to assess their sources. It might therefore be expected that the sceptic, who could be characterised as being the least likely to attend to advice given by others, would have largely ignored the guidance given. This was not true; indeed, it appeared that this aspect of the intervention was as effective as for the sceptic as it was for two of the moderate multiplists and may have been even more effective if he had allocated additional time to engaging in critical evaluation.

8 Discussion

8.1 Introduction

This research aimed to identify the challenges that Science FY students with differing justification beliefs face when constructing arguments from multiple documents. It also aimed to develop and evaluate an intervention which supports these students in constructing arguments from multiple documents. Consequently, this concluding chapter summarises the extent to which this aim was met by discussing the key findings, considering how these findings relate to the literature, and indicating the contributions made by this research. The limitations of this research are then considered, and recommendations for future research are given. The final section describes the implications of this research for future practice.

8.2 Key Findings in Relation to the Research Questions

This research addressed four research questions:

RQ1 What challenges do Science FY students at the researched university face when constructing arguments from multiple documents?

RQ2 What are the justification beliefs of these students?

RQ3 What are the design principles of an intervention that seeks to support these students in constructing arguments from multiple documents?

RQ4 What are students' experience of and outcomes from this intervention?

The literature review (Chapter 3) indicated that justification beliefs affect multiple document comprehension and engagement in argumentation and may interact with prior knowledge and topic interest. This section begins by discussing the justification beliefs of Science FY students, which addresses *RQ2***.** Students' justification beliefs were expected to affect the challenges they faced when trying

to construct an argument from multiple documents **(***RQ1***);** these challenges were explored during Study 1. Key findings from the thematic framework developed through this research are discussed.

Participants with differing justification beliefs were expected to have different experiences of and outcomes from the intervention. These experiences and outcomes were primarily explored in Study 2. Since these findings offer an evaluation of the key design principles, the findings in relation to *RQ3* are discussed after discussion of participants' experiences and outcomes.

8.2.1 What are the Justification Beliefs of Science FY Students?

Students' justification beliefs were measured using the Justification for Knowing Questionnaire (JFK-Q), which was developed by Ferguson et al. (2013) in relation to justification beliefs in the natural sciences. Using 14 items, the JFK-Q measures beliefs in personal justification, beliefs in justification by authority and beliefs in justification by multiple sources by asking students to rate how strongly they agree with statements such as "What is a fact in the natural sciences depends on one's personal views", "If a scientist says that something is a fact, then I believe it", and "I can never be sure about a claim in the natural sciences until I have checked it with at least one other source".

The researched students were found to be relatively homogenous with respect to the strength of their beliefs in justification by multiple sources but were diverse with respect to their beliefs in personal justification and justification by authority. Participants were sorted into groups with others who had similar patterns of justification beliefs by means of Ward's method of cluster analysis (Everitt, 2011). In both studies, the largest clusters contained those students who initially had similar strength beliefs in personal justification and justification by authority. These participants were named the multiplists, a term derived from Kuhn's (1991) model of informal reasoning, which describes those who believe all viewpoints to be equally valid. Given that scientists consider evidence collected by scientific means as having greater validity than personal experiences, having multiplist beliefs was expected to be maladaptive when constructing scientific

arguments. Ten of the 25 participants (40%) in the Study 1 cohort and 17 of the 30 participants (57%) in the Study 2 cohort were multiplists.

Most of the remaining participants in each cohort were dogmatists, though a small proportion of each cohort were sceptics. These terms, which originated in Chandler et al.'s (1990) work on the dogmatism-scepticism axis, were used by Greene et al. (2008) to describe the contrasting ways in which students may respond when faced with conflicting perspectives. Dogmatists look to those with expertise to decide which claims are likely to be true, whereas those who are sceptics disregard such views and instead rely on personal experiences to justify their knowledge claims. Here, the dogmatists were distinguished from the sceptics by comparing the strengths of their beliefs in personal justification and justification by authority. Those with relatively weak beliefs in personal justification and relatively strong beliefs in authority were named the dogmatists and those with relatively strong beliefs in personal justification and relatively weak beliefs in authority were sceptics. In Study 1, just over a third (36%) were dogmatists and just under a quarter (24%) were sceptics. The proportion of dogmatists in the Study 2 cohort was lower, with just over a quarter (27%) having such beliefs. The Study 2 cohort also included a lower proportion of sceptics, with just 17% of participants being sorted into this cluster. Having strong beliefs in personal justification is likely to be maladaptive when working with multiple documents (Strømsø et al., 2011) because students with such beliefs tend to attend less to the content of documents. While being a dogmatist is thought to be adaptive when students are learning a new discipline, an over-reliance on expert views may limit knowledge gains (Bråten et al., 2008, 2009).

8.2.2 What Challenges do Science FY Students Face When Constructing Arguments From Multiple Documents?

Study 1 provided an opportunity for Science FY students to describe their experiences of constructing an argument from multiple documents. The experiences of a subset of the Study 1 participants were captured through recordings of individual tutorials, within a reflective piece of writing, and during retrospective interviews. The heterogeneity of justification beliefs within the Study

1 cohort was reflected in the participants chosen for interview. Two of the interviewees were from the multiplist cluster, two were from dogmatist cluster and two were from the sceptics cluster. Although it was expected that the interviewees would face different challenges due to their different justification beliefs, caution is needed in generalising from these participants due to the small sample sizes. Consequently, differences between those with different justification beliefs are only noted when it is appropriate to do so.

The Cognitive Challenges theme, which dominated the thematic framework developed within this research, expressed the difficulties that many of the interviewees faced when trying to select, read, critically evaluate, and synthesise documents. The reviewed literature rightly predicted that justification beliefs would interact with the cognitive habits needed to comprehend multiple documents and engage in argumentation (Bråten & Strømsø, 2010; Gil et al., 2010). Within this research, the accounts of the interviewed multiplists have provided a vivid illustration of how relatively equal strength beliefs in personal justification and beliefs in authority are maladaptive, particularly when students have not yet become skilled readers. For instance, Mollie's overly broad search strategy hindered her ability to create a coherent mental representation and left her feeling her overwhelmed by the sheer number of sources she was trying to read. Given that Mollie aligned closely with the Affectively Engaged stance described within the CAEM (List & Alexander, 2017), her experiences were not unexpected. Another of the interviewed multiplists, Matthew, also described having faced multiple cognitive challenges. After spending considerable time trying to synthesise information from multiple documents, Matthew was reluctant to engage in critical evaluation because he was concerned about not having enough time to find replacement sources.

Participants such as Mollie and Matthew who experienced cognitive challenges often experienced negative affect. Many described how they had felt frustrated, worried, or uncertain. As such, these participants often needed reassurance. Furthermore, these participants often adopted maladaptive coping strategies, with procrastination being the most common. It is therefore concluded that the negative affect experienced by students who are yet to develop effective

cognitive sourcing and corroboration behaviours is likely to hinder multiple document comprehension.

The reviewed literature also indicated that prior educational experiences may determine whether students can meet the cognitive challenges associated with constructing an argument. The findings from this research support the assumption that secondary school students rarely engage in argumentation during science classes (Duschl & Osborne, 2002). Even when students had engaged in argumentation in other disciplines, they often lacked confidence in constructing scientific arguments.

A novel finding from this research is that lack of experience in working with multiple documents leaves students unaware of how long is needed to select, read, critically evaluate, and synthesise information from multiple documents.

8.2.3 What were Students' Experiences of and Outcomes from the Intervention?

The designed intervention aimed to address the challenges that Science FY students face when trying to construct arguments from multiple documents. Changes to participants' justification beliefs, which were measured using the JFK-Q, provided one of the outcome measures. Participants' experiences were gathered during individual tutorials, a piece of reflective writing, and a retrospective interview.

A key finding from this research is that participants' justification beliefs became more adaptive after the intervention. In both studies, participants' beliefs in personal justification were weaker and their beliefs in justification by multiple sources were stronger in the post-intervention measure. This manifested as a significant decrease in beliefs in Personal justification in Study 1. The strengthening of beliefs in multiple sources was statistically significant for the Study 2 participants, and as such it contributed to a statistically significant change in these participants' justification beliefs. These findings are consistent with those of Ferguson et al. (2013), who found that there is a tendency for beliefs in justification by multiple sources to strengthen and beliefs in personal justification to weaken when students work with multiple documents that have conflicting perspectives. Rosman et al. (2019) explained that reading multiple documents can help students to recognise that not all perspectives are equally valid, resulting in a weakening of beliefs in

personal justification. This was evident in the present study. For example, Marc, described how he had come to realise that sources differ in their credibility. Others also talked about how they had learnt that not all sources should be trusted.

Schoor (2023) recently proposed that reliance on personal justification is associated with a lack of trust in scientific institutions and those with expertise in science. Comments made by some of the interviewed multiplists and sceptics in the present study suggested that they did not know how to determine who to trust. Learning how to assess author expertise lessened their reliance on personal justification and, in some cases, transformed their understanding of what constitutes valid evidence in science. Others also reduced their reliance on personal justification through developing their sourcing behaviours. It can therefore be concluded that this research has established that teaching students how to assess whether a source is credible and trustworthy can lessen their reliance on personal justification beliefs.

It must, however, be acknowledged that not all participants had lower scores for beliefs in personal justification after the intervention. Given that Wiley and Voss (1996), Le Bigot and Rouet (2007), and Bråten and Strømsø (2011) have all found that constructing arguments from multiple documents develops conceptual understanding, one possible explanation is that Science FY students became more knowledgeable when completing the Making an Argument assignment. Hence, the stronger beliefs in personal justification reported by some participants in the post-intervention measure may indicate that having worked with multiple documents to construct an argument these participants felt able to draw upon their own knowledge to justify their knowledge claims.

8.2.4 What are the Key Design Principles of an Intervention that Supports Science FY Students in Constructing Arguments from Multiple Documents?

Participants' experiences of the intervention have enabled the effectiveness of the design principles that underpinned the intervention to be explored. These design principles were originally founded in the literature reviewed in Chapter 3. Their implementation was described in Chapter 5 and their effectiveness was initially

explored in Study 1 (Chapter 6). Revisions to the design principles and their implementation were made in light of the findings from Study 1; these revisions are described in Chapter 7. Chapter 7 also discusses how Study 2 provided an opportunity to explore the effectiveness of the revised principles. The findings from Study 2 in relation to the key design principles are discussed here. These findings have allowed conclusions to be reached about whether each design principle should be retained, refined, or removed in any future iterations of the intervention. These conclusions provide the basis of the four recommendations for future practice.

Design Principle 1

Design Principle 1, which was informed by Kienhues (2016) and Sandoval (2016), was concerned with providing students with opportunities to develop their disciplinary knowledge and domain-specific beliefs. von der Mühlen et al., (2019) proposed that reading scientific publications as a means of developing disciplinary knowledge is likely to promote adaptive domain-specific beliefs because students will encounter claims supported by evidence. It should, however, be noted that the findings from this research suggest that students with limited pre-existing knowledge who attempt to develop their scientific knowledge through reading journal articles are likely to face cognitive challenges and experience negative affect. For instance, in Study 1, participants who attempted to become more knowledgeable by reading journal articles struggled with the complexities of documents written for those with greater subject expertise and found it challenging to synthesise conflicting information. Having experienced cognitive challenges, these participants experienced negative affect, which resulted in the adoption of maladaptive coping strategies.

Design Principle 1 was subsequently revised to become focused on providing students with clear guidance about how to select and use appropriate sources of information, such as textbooks and review articles, when learning new disciplinary knowledge. The findings from Study 2 indicate that participants paid close attention to the guidance given about how to evaluate source information and that the multiplists acted upon this guidance by changing the sources they selected. It

has therefore been concluded that it is worth retaining the revised version of this design principle in any future iterations of the intervention.

Design Principle 2

Design Principle 2 sought to encourage students to select socio-scientific issues of interest to them. This design principle was somewhat controversial because the role of topic interest in multiple document comprehension is contested. Having high levels of topic interest which are not matched by prior knowledge may be problematic because such students are more likely to rely on those with expertise to justify their knowledge claims (Bråten et al., 2009), which may limit their knowledge gains (Bråten et al., 2008). Furthermore, Brandmo and Bråten (2018) contended that students with strong personal justification beliefs may be less interested in complex or controversial topics than those with other justification beliefs. Their expectations did not match the findings of the present study, where all participants selected socio-scientific issues that matched their personal, career, or academic interests.

The decision to adopt a design principle that promotes topic interest reflects the context in which the researched students are working with multiple documents. The Constructing an Argument assignment requires these students to work with multiple documents over an extended period while also studying three content-driven modules. This research has taken the position that allowing students who are working with multiple documents under similar conditions to select topics that match their interests is likely to sustain their efforts.

Design Principle 3

Design Principle 3 was about teaching students to construct arguments. This design principle was influenced by Duschl and Osborne (2002), who proposed that Toulmin's (1958) model of practical arguments provides a useful framework for scientific arguments. While this may be true for more experienced scientists, the experiences of the Study 1 participants indicate that students who are relatively new to the discipline struggle to comprehend some of the components. Consequently, attempts to engage Science FY students with this framework and

with arguments written by others failed to achieve their intended aim of teaching students how to construct arguments from these components. Furthermore, it was also evident that the researched participants found it challenging to incorporate information from multiple sources when attempting to construct their own arguments. These findings motivated changes to this design principle ahead of Study 2. The revised principle sought to teach students about the components of arguments and how to find and use information from multiple sources to construct arguments from these components.

The experiences of the Study 2 participants further support the conclusion reached after Study 1 that teaching students about the components of arguments is not sufficient; it is only when students construct their own arguments that they fully grasp how to structure an argument. This is not to say that the resources used to teach students about arguments were not effective; indeed, many of the Study 2 participants commented on how the examples given and the approach used to draw attention to different components had proved valuable when constructing their own arguments. Any future iterations of the intervention should continue to use these resources and approaches to teach students about the structure of arguments ahead of providing them with the opportunity to construct their own argument.

Design Principle 4

Design Principle 4 was not effective in its original form. This design principle was somewhat ambitious in its aim of equipping students with the cognitive habits needed to engage with multiple documents. The documents model, which provides the foundations for the two models of MDC adopted within this research, assumes that those working with multiple documents have effective sourcing and corroboration behaviours. Design Principle 4 sought to equip students with these cognitive habits by providing them with checklists and templates that prompted them to record key details and engage in critical evaluation.

Study 2 participants valued having prompts to guide their reading, though there was some evidence that participants' justification beliefs affected the extent to which they attended to and adopted the guidance given. Those with strong

justification by authority beliefs and those with similar strength beliefs in personal justification and authority often spoke of how they had valued the checklists and templates provided during the workshops.

The researched students were, however, rightly concerned about whether the notes made in response to these prompts would be of use when constructing an argument. Nonetheless, the findings from this research indicate that Science FY students benefit from being given guidance about how to locate evidence in support of claims, and information that can be used as warrants and their backings when reading multiple documents. It has also been concluded that students studying science at university should also be taught how to critically evaluate scientific research and why doing so will enable them to incorporate qualifiers and rebuttals into their claims.

Design Principle 5

The fifth design principle acknowledged the critical role of motivation in validation of conflicting information (Richter & Maier, 2017). In this research, students were given timely, positive feedback during the workshop activities and through an individual tutorial. All interviewed participants described wholly positive outcomes from being given positive feedback. This design principle should be retained because this research has provided evidence that providing positive, timely feedback motivates and encourages students to allocate cognitive resources even when faced with other demands on their time and efforts.

8.2.5 Recommendations for Future Practice

Science FY students do not arrive at university knowing how to construct arguments from multiple documents. The findings from this research suggest that Science FY students with differing justification beliefs can construct arguments from multiple documents if given appropriate support. This support should be founded on four design principles:

DP1: Students who are working with multiple documents over an extended period should be given the opportunity to select a socio-scientific issue that interests them.

DP2: Those with similar strength beliefs in personal justification and justification by authority should be given clear guidance about how to identify and use appropriate sources of information when developing their disciplinary knowledge by being guided towards textbooks and review articles.

DP3: Students should be taught how to select sources that provide what scientists constitute to be valid evidence in support of claims. This can be achieved by providing resources such as checklists and templates that prompt them to consider how data were collected, analysed, and interpreted. Science FY students should also be taught how to engage in critical evaluation and why doing so will enable them to incorporate qualifiers and rebuttals into their claims when making arguments. Science FY students who lack prior experience of working with multiple documents should be given clear guidance about the time needed to select, read, critically evaluate, and synthesise information from multiple documents.

To fully grasp the roles played by qualifiers and rebuttals, and other components of arguments such as the warrant or backings, students must be given opportunities to construct their own arguments and should be given timely, formative feedback about their initial attempts.

DP4: Those working with Science FY students need to recognise that engaging with multiple documents is cognitively challenging and can lead to negative affect. Students need to be given opportunities to seek clarification in a timely manner and be given positive feedback, particularly when they are juggling multiple demands on their time. Doing so should motivate students to persist in effortful cognitive processing and ameliorate any unhelpful consequences of having experienced negative affect.

8.3 Contributions

I have designed an effective intervention which supports students with differing justification beliefs in constructing arguments from multiple documents. The

findings from this research enable contributions to be made to knowledge, methodology, and professional practice.

8.3.1 Contributions to Knowledge

This research drew upon literature at the intersections between multiple document comprehension, argumentation, and justification beliefs. Regarding the latter, the reviewed literature provided a hierarchy of justification beliefs, with beliefs in justification by authority being most adaptive when students lack prior knowledge (Bråten et al., 2008) and beliefs in personal justification being least adaptive (Strømsø et al., 2011). Based on the findings from this research, it is proposed that those with similar strength beliefs in personal justification and justification by authority should be added to this hierarchy, where they would occupy the lowest position. This is because students with such beliefs often try to accommodate all perspectives, which proves problematic when trying to corroborate information.

In addition, this research has provided some evidence that the adoption of effective cognitive habits may be associated with the adoption of more adaptive justification beliefs. Those with relatively strong personal justification beliefs and relatively weak beliefs in justification by authority who learnt how to assess an author's credibility and how to extract evidence in support of claims from documents lessened their reliance on personal experiences or placed more trust in those with expertise. These findings suggest that teaching students how to judge expertise, assess the validity of alternative perspectives, and use information from documents as evidence in support of claims has the potential to bring about adaptive changes to justification beliefs.

This research has also found that students who lack experience in working with multiple documents rarely allocate enough time to find, read, and evaluate information. It has also been concluded that procrastination is often a maladaptive response to experiencing negative affect. The provision of timely feedback through the individual tutorial was crucial in re-engaging students. The positive outcomes from and experiences of being given positive feedback support Richter and Maier's (2017) position that motivation plays a critical role in encouraging students to allocate cognitive resources. Moreover, it is likely that participants' high levels of

interest in the socio-scientific issues sustained their efforts over an extended period.

8.3.2 Methodological Contributions

This research adopted the Justification for Knowing Questionnaire (JFK-Q) as a means of measuring the researched students' justification beliefs before and after the intervention activities had taken place. In doing so, this research aligned itself with Greene et al. (2008), who justify the use of Likert scales as a means of sorting heterogenous samples. In this research, the JFK-Q was a valuable tool for distinguishing between participants with different justification beliefs. Indeed, the use of cluster analysis to systematically select participants with different characteristics was a key methodological contribution of this research.

Having conducted this research within a university, it is proposed that the JFK-Q should be amended when being used in similar educational settings. The item that refers to textbooks should be edited to instead refer to research publications. This is because university students are expected to engage with primary rather than secondary sources when providing evidence in support of claims. Those items that currently refer to 'sources' or 'information sources' should also be amended to refer to research publications. This is because the findings of this research indicate that some students initially viewed websites aimed at the public as being as credible as peer-reviewed scholarly sources. It is also proposed that some of the items which currently refer to scientists should instead refer to practitioners. This is because many university students are pursuing professional degrees in which the expertise of a practitioner is as valid as that of a scientist. Although a relatively minor change, references to science teachers should be changed to tutor when administering the questionnaire to university students.

This research has demonstrated it is possible research multiple document comprehension, argumentation, and justification beliefs within the educational context in which learning takes place. Doing so within the present study necessitated the use of qualitative methods alongside self-report measures, something which Bråten et al (2009) urged researchers to do in order to better enable them to explore epistemic beliefs in action. Though gathering data from

transcripts lacked the ease and efficiency of measures used in experimental studies, the resulting data provided valuable insights. For example, in this study, remarks made by students during their individual tutorials indicated that students' beliefs in personal justification became weaker once they had learnt how to assess whether sources were credible and trustworthy.

The findings of this research make a further methodological contribution. In their meta-analysis of epistemic cognition interventions, Cartiff et al. (2021) concluded that interventions lasting more than a month were unlikely to result in statistically significant results, though they did note that such studies typically measured the effectiveness of the outcome using measures of conceptual knowledge or academic achievement, not epistemic cognition. In this research, the post-intervention measures in Study 2 were completed two months after the intervention began. Across all scales of the JFK-Q, students demonstrated more adaptive justification beliefs, becoming less reliant on personal justification beliefs and more reliant on justification by multiple sources. Hence, this research has demonstrated that when pre- and post-test measures of justification beliefs are used to measure the effectiveness of an epistemic cognition intervention, it is possible to demonstrate that longer-term interventions are associated with positive changes to justification beliefs.

8.3.3 Contributions to Professional Practice

This research has contributed to my professional practice. For instance, my understanding of argument structure has changed. I originally considered the rebuttal to be an aspect of a counterclaim. When returning to Toulmin's (1958) framework after Study 1, I recognised that he conceived the rebuttal as being a means of acknowledging exceptions or objections to the claim. This shift in my understanding has brought about changes to the Making an Argument assignment, whereby students are no longer asked to include a counterclaim and rebuttal.

More importantly, I am now able to anticipate the challenges that undergraduate science students may face when asked to find, read, evaluate, and synthesise multiple documents. Consequently, I have adapted the intervention activities developed through this research to create two, two-hour workshops that

support second year undergraduate science students in writing literature reviews that identify different perspectives about a scientific issue and reach a conclusion about which perspective is currently the consensus view. These workshops feature activities and resources that were designed and evaluated during this research.

8.4 Limitations and Recommendations for Further Research

The workshop, tutorial, and assessment activities associated with the intervention were embedded into a compulsory module of the Science FY course. As such, it was not appropriate to have a control group and therefore a single-group pre-test post-test design was adopted. This method can face threats to its internal validity, such as history effects, maturation effects, instrument sensitivity, and regression to the mean. Regarding the former, the designed intervention activities in this research were delivered alongside a session by the university's library team about how to find sources using search engines and databases. While it is possible that this session could have contributed to changes in their justification beliefs, there was little evidence that students who engaged with this session had a better understanding of what constitutes an appropriate scientific source.

Although justification beliefs are thought to change more slowly than other aspects of epistemic cognition (Muis & Duffy, 2013), maturation effects may have contributed to the reported changes to justification beliefs because the research was conducted over several weeks. It is also possible that completing the pre-intervention JFK-Q enabled students to gain awareness of the different ways in which they could justify their knowledge and consequently led some students to change how they justified what they believe to be true. A further limitation is that some students' scores on the pre-intervention JFK-Q were at the extremes of the scale, leaving them little or no room to demonstrate positive changes to their justification beliefs after the intervention.

The relatively small cohorts studied within this research lacked statistical power. Additionally, due to the use of Likert scales, it was necessary to use non-parametric statistical tests, which are not as powerful as those that assume the Normal Distribution. It is therefore possible that the intervention had small effects on justification beliefs, but the lack of power resulted in Type II errors. At the time

this research was conducted, I had limited access to undergraduate science students. It would now be possible for future research to be conducted using larger cohorts.

Design-based research was chosen because it allows for progressive refinement, but the iterative nature of this methodology can limit empirical control. Embedding the intervention along other teaching, learning, assessment, and feedback activities, and making changes between Studies 1 and 2 may have limited the empirical control, though these changes are likely to have led to better outcomes for the participants (Sandoval & Bell, 2004; Bakker & van Eerde, 2015). Unlike many DBR projects, there was no need to hand over control of the implementation of the educational intervention to a third party, though this meant that it was necessary to deal with the ethical issues associated with being a practitioner-researcher. These included minimising power relations and having limited analytical distance. Regarding the former, efforts were made to always refer to myself as being a researcher when administering the JFK-Q and conducting the interviews. Nonetheless, some of the interviewees raised matters that were outside the scope of this research, suggesting that they were viewing me as their teacher. It was also challenging to always maintain an appropriate analytical distance, particularly when it was evident that one of the participants was experiencing emotional distress.

8.5 Conclusions

Science FY students at the researched university differ in how they justify what they believe to be true. Those with similar strength beliefs in personal justification and justification by authority, who are termed the multiplists, are likely to face numerous cognitive challenges when trying to construct scientific arguments from multiple documents. Having not yet developed effective sourcing and corroboration strategies, they may become overwhelmed when reading documents with conflicting perspectives. Although multiplists are likely to seek, and act upon, guidance about how to manage multiple documents, their initial progress may be hindered by their need for reassurance.

The designed intervention was successful in bringing about adaptive changes to Science FY students' justification beliefs. Reliance on justification by personal justification decreased, particularly when students paid attention to guidance given about how to determine which sources are credible and trustworthy. Participants' beliefs in justification by multiple sources increased.

The design principles that underpinned the designed intervention were informed by the literature and revised in the light of the experiences of those who experienced its implementation. It has been concluded that when university students who are studying science subjects are required to work with multiple documents over an extended period, four principles should guide the support they are given. Firstly, students should select issues that interest them. Secondly, they should be given clear guidance about how to identify and use appropriate sources of information when developing their disciplinary knowledge by being guided towards textbooks and review articles, particularly if they are multiplists. Thirdly, they should be taught what constitutes valid evidence in science, given resources that enable them to record salient information from multiple documents, taught about the value of engaging in critical evaluation of scientific research when constructing an argument, and given opportunities to construct their own arguments. Finally, those who need clarification or reassurance should be able to access positive feedback in a timely manner.

This research has contributed to our knowledge, made methodological contributions, and has informed professional practice. Considering the findings from this research, it has been proposed that having similar strength beliefs in personal justification and justification by authority should be considered as being maladaptive, even more so than having strong beliefs in personal justification. It has also been proposed that developing students' cognitive capabilities through teaching them more effective sourcing behaviours has the potential to bring about adaptive changes to justification beliefs. Furthermore, this research has contributed to the debate about the role of topic interest; it has been concluded that, along with the provision of positive feedback, promoting high levels of topic interest is likely to sustain students' efforts and motivate them to allocate cognitive resources even when they are facing multiple demands on their time.

The key methodological contribution of this research is that cluster analysis offers a highly effective means of ensuring the heterogeneity of a cohort is reflected in the participants selected for interview. This research has also demonstrated that embedding research about argumentation, multiple document comprehension, and justification beliefs into ecologically valid settings provides valuable insights into the experiences of learners that aid interpretation and enable the findings to make meaningful contributions to future educational practice.

Having designed, implemented, and evaluated the educational intervention at the heart of this research, I have developed my understanding of the challenges that university students face when working with multiple documents. Consequently, this research has contributed to my professional practice because I am now able to develop activities and resources that support university students through the cognitive and affective challenges of working with multiple documents.

References

Acosta-Gonzaga, E., & Ramirez-Arellano, A. (2022). Scaffolding matters? Investigating its role in motivation, engagement and learning achievements in higher education. *Sustainability, 14*(20), 13419. https://doi.org/10.3390/su142013419

Andriessen, J., & Baker, M. (2014) Arguing to learn. In: R. K. Sawyer (Ed.), *The Cambridge handbook of the Learning Sciences.* (2nd edition). (pp.439-460). Cambridge University Press.

Badley, G. (2003). The crisis in educational research: a pragmatic approach. *European Educational Research Journal, 2*(2), 296-308. https://doi.org/10.2304/eerj.2003.2.2.7

Bakker, A. (2018). *Design research in education: A practical guide for early career researchers.* Routledge.

Bakker, A., & van Eerde, D. (2015). An introduction to design-based research with an example from statistics education. In C. Knipping, & N. Presmeg. (Eds.) *Approaches to qualitative research in mathematics education: Examples of methodology and methods.* (pp. 429-466). Springer.

Barab, S. (2006). Design-based research: A methodological toolkit for the learning scientist. In K. Sawyer (Ed.) *The Cambridge handbook of the Learning Sciences.* (2nd edition). (pp. 153-169). Cambridge University Press.

Belland, B. R., Kim, C., & Hannafin, M. J. (2013). A framework for designing scaffolds that improve motivation and cognition. *Educational Psychologist, 48*(4), 243-270. https://doi.org/10.1080/00461520.2013.838920

Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology*, *38*(3), 180-195. https://doi.org/10.1016/j.cedpsych.2013.03.003

Brandmo, C., & Bråten, I. (2018). Investigating relations between beliefs about justification for knowing, interest, and knowledge across two socio-scientific

topics. *Learning and Individual Differences*, *62*, 89-97. https://doi.org/10.1016/j.lindif.2018.01.010

Bråten, I. (2016). Epistemic cognition interventions: Issues, challenges, and directions. In J. A. Greene, W. Sandoval, W., & I. Bråten (Eds.) *Handbook of epistemic cognition*. (pp. 372-384). Routledge.

Bråten, I., Braasch, J. L., Strømsø, H. I., & Ferguson, L. E. (2015). Establishing trustworthiness when students read multiple documents containing conflicting scientific evidence. *Reading Psychology, 36*(4), 315-349. https://doi.org/10.1080/02702711.2013.864362

Bråten, I., Brante, E. W., & Strømsø, H. I. (2018). What really matters: The role of behavioural engagement in multiple document literacy tasks. *Journal of Research in Reading*, *41*(4), 680-699. https://doi.org/10.1111/1467-9817.12247

Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (2014). Students working with multiple conflicting documents on a scientific issue: Relations between epistemic cognition while reading and sourcing and argumentation in essays. *British Journal of Educational Psychology*, *84*(1), 58-85. https://doi.org/10.1111/bjep.12005

Bråten, I., Gil, L., Strømsø, H. I., & Vidal-Abarca, E. (2009). Personal epistemology across cultures: Exploring Norwegian and Spanish university students' epistemic beliefs about climate change. *Social Psychology of Education, 12*, 529-560. https://doi.org/10.1007/s11218-009-9097-z

Bråten, I., & Strømsø, H. I. (2010). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes*, *47*(1), 1-31. https://doi.org/10.1080/01638530902959646

Bråten, I., Strømsø, H. I., & Samuelstuen, M. S. (2008). Are sophisticated students always better? The role of topic-specific personal epistemology in the understanding of multiple expository texts. *Contemporary Educational Psychology*, *33*(4), 814-840. https://doi.org/10.1016/j.cedpsych.2008.02.001

British Educational Research Association (2018) *Ethical guidelines for educational research* (4th edition). BERA.

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J. F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. W. van den Broek (Eds). *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 209-233). Lawrence Erlbaum Associates.

Britt, M. A., & Rouet, J.-F. (2012) Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby, & M. J. Lawson. (Eds.) *Enhancing the quality of learning dispositions, instruction, and learning processes.* (pp. 276-314). Cambridge University Press.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2* (2), 141-178. https://doi.org/10.1207/s15327809jls0202_2

Cartiff, B. M., Duke, R. F., & Greene, J. A. (2021). The effect of epistemic cognition interventions on academic achievement: A meta-analysis. *Journal of Educational Psychology, 113* (3), 477. https://doi.org/10.1037/edu0000490

Chan, J., Fancourt, N., & Guilfoyle, L. (2021). Argumentation in religious education in England: an analysis of locally agreed syllabuses. *British Journal of Religious Education*. *43* (4), 458-471. https://doi.org/10.1080/01416200.2020.1734916

Chandler, M., Boyes, M., & Ball, L. (1990). Relativism and stations of epistemic doubt. *Journal of Experimental Child Psychology*. *50* (3), 370-395. https://doi.org/10.1016/0022-0965(90)90076-K

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher. 32* (1), 9-13. https://doi.org/10.3102/0013189X032001009

Cochran-Smith, M., & Lytle, S. L. (2007). Everything's ethics: Practitioner inquiry and university culture. In A. Campbell, & S. Groundwater-Smith (Eds.) *An ethical*

approach to practitioner research: Dealing with issues and dilemmas in action research. (pp. 40-57) Routledge.

Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.) *New directions in educational technology* (pp. 123-130). Springer.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, 13(1), 15–42. https://doi.org/10.1207/s15327809jls1301_2

Cranmer, G. (Ed.) (2017). *The SAGE encyclopaedia of communication research methods.* Vol. 4. Sage Publications.

Creswell, J. W. (2009) *Research design: Qualitative, quantitative, and mixed methods approaches.* (3rd edition). Sage Publications.

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. Sage Publications.

Department for Education. (2013). Science programmes of study: Key Stage 3 National Curriculum in England. https://www.gov.uk/government/publications/national-curriculum-in-england-science-programmes-of-study.

Department for Education (2014). Science programmes of study: Key stage 4 National Curriculum in England. https://www.gov.uk/government/publications/national-curriculum-in-england-science-programmes-of-study.

Derrick, B., & White, P. (2017). Comparing two samples from an individual Likert question. *International Journal of Mathematics and Statistics, 18*(3), 1-13.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*(3), 287-312. https://doi.org/10.1002/(SICI)1098-237X(200005)84:3%3C287::AID-SCE1%3E3.0.CO;2-A

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education*, Studies in Science Education, 38*(1), 39-72. https://doi.org/10.1080/03057260208560187

Edelson, D. C. (2002) Design research: What we learn when we engage in design. *Journal of the Learning Sciences, 11*(1), 105-121. https://doi.org/10.1207/S15327809JLS1101_4

Erduran, S., Guilfoyle, L., & Park, W. (2022). Science and Religious Education teachers' views of argumentation and its teaching. *Research in Science Education, 52*, 655-673. https://doi.org/10.1007/s11165-020-09966-2

Everitt, B. (2011). *Cluster analysis* (5th ed.). Wiley-Blackwell.

Ferguson, L. E., & Bråten, I. (2013). Student profiles of knowledge and epistemic beliefs: Changes and relations to multiple-text comprehension. *Learning and Instruction*, *25*, 49-61. https://doi.org/10.1016/j.learninstruc.2012.11.003

Ferguson, L. E., Bråten, I., & Strømsø, H. I. (2012). Epistemic cognition when students read multiple documents containing conflicting scientific evidence: A think-aloud study. *Learning and Instruction*, *22*(2), 103-120. https://doi.org/10.1016/j.learninstruc.2011.08.002

Ferguson, L. E., Bråten, I., Strømsø, H. I., & Anmarkrud, Ø. (2013). Epistemic beliefs and comprehension in the context of reading multiple documents: Examining the role of conflict. *International Journal of Educational Research*, 62, 100-114. https://doi.org/10.1016/j.ijer.2013.07.001

Fraser, D. M. (1997). Ethical dilemmas and practical problems for the practitioner researcher. *Educational Action Research, 5*(1), 161-171. https://doi.org/10.1080/09650799700200014

General Data Protection Regulation. (2018). *Guide to the General Data Protection Regulation.* [online] GOV.UK. https://www.gov.uk/government/publications/guide-to-the-general-data-protection-regulation

Gibbs, G. (1988). Learning by doing: A guide to teaching and learning methods. *Further Education Unit.*

Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010). Understanding and integrating multiple science texts: Summary tasks are sometimes better than argument tasks. *Reading Psychology, 31*(1), 30-68. https://doi.org/10.1080/02702710902733600

Goldsmith, L. J. (2021). Using Framework analysis in applied qualitative research. *The Qualitative Report*, 26(6), 2061-2076. https://doi.org/10.46743/2160-3715/2021.5011

Green, H. J., & Hood, M. (2013). Significance of epistemological beliefs for teaching and learning psychology: A review. *Psychology Learning & Teaching*, *12*(2), 168-178. https://doi.org/10.2304/plat.2013.12.2.168

Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modelling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist*, *43*(3), 142-160. https://doi.org/10.1080/00461520802178458

Greene, J. A., Sandoval, W. A., & Bråten, I. (2016). Reflections and future directions. In J. A. Greene, W. Sandoval, & I. Bråten (Eds.) *Handbook of epistemic cognition*. (pp.495-510). Routledge.

Guilfoyle, L., Erduran, S., & Park, W. (2021). An investigation into secondary teachers' views of argumentation in science and religious education. *Journal of Beliefs & Values*, *42*(2), 190-204. https://doi.org/10.1080/13617672.2020.1805925

Guilfoyle, L., Hillier, J., & Fancourt, N. (2021). Students' argumentation in the contexts of science, religious education, and interdisciplinary science-religious education scenarios. *Research in Science & Technological Education,* 1-18. https://doi.org/10.1080/02635143.2021.1947223

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (7th edition). (2014). *Multivariate data analysis.* Pearson Education Ltd.

Hatano, G., & Inagaki, K. (1984). Two courses of expertise. *Research and Clinical Center for Child Development, 6,* 27-36. http://hdl.handle.net/2115/25206

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, *67*(1), 88-140. https://doi.org/10.3102/00346543067001088

Hotulainen, R., & Telivuo, J. (2015). Epistemological beliefs and scientific reasoning in Finnish academic upper secondary education, *Kasvatus & Aika, 9*(1), 92-106. University of Helsinki Open Repository. http://hdl.handle.net/10138/162963

Iordanou, K., Kendeou, P., & Beker, K. (2016). Argumentative reasoning. In J. A. Greene, W. Sandoval, W., & I. Bråten (Eds.) *Handbook of epistemic cognition*. (pp. 51-65). Routledge.

Kendeou, P., Braasch, J. L. G., & Bråten, I. (2016). Optimizing conditions for learning: Situating refutations in epistemic cognition. *The Journal of Experimental Education*, *84 (2),* 245-263. https://doi.org/10.1080/00220973.2015.1027806

Kienhues, D., Bromme, R., & Stahl, E. (2008). Changing epistemological beliefs: The unexpected impact of a short-term intervention. *British Journal of Educational Psychology*, *78*(4), 545-565. https://doi.org/10.1348/000709907X268589

Kienhues, D., Ferguson, L., & Stahl, E. (2016). Diverging information and epistemic change. In J. A. Greene, W. Sandoval, & I. Bråten (Eds.) *Handbook of epistemic cognition*. (pp. 330-342). Routledge.

Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education, 77*(3), 319–337. https://doi.org/10.1002/sce.3730770306

Lewis, J., & Speers, T. (2003). Misleading media reporting? The MMR story. *Nature Reviews Immunology*, 3, 913–918. https://doi.org/10.1038/nri1228

Le Bigot, L., & Rouet, J.-F. (2007). The impact of presentation format, task assignment, and prior knowledge on students' comprehension of multiple online documents. *Journal of Literacy Research*, *39*(4), 445-470. https://doi.org/10.1080/10862960701675317

List, A., & Alexander, P. A. (2017). Cognitive affective engagement model of multiple source use. *Educational Psychologist*, *52*(3), 182-199. https://doi.org/10.1080/00461520.2017.1329014

Litman, C., Marple, S., Greenleaf, C., Charney-Sirott, I., Bolz, M. J., Richardson, L. K., Hall, A., George, M., & Goldman, S. R. (2017). Text-based argumentation with multiple sources: A descriptive study of opportunity to learn in secondary English language arts, history, and science. *Journal of the Learning Sciences*, *26*(1), 79-130. https://doi.org/10.1080/10508406.2016.1256809

Lunt, N., & Fouché, C. (2010). Practitioner research, ethics and research governance. *Ethics and Social Welfare*, 4(3), 219-235. https://doi.org/10.1080/17496535.2010.516117

McArdle, K., Birchley, J., Bruce, J., Hurrell, A., Paterson, S., & Stephen, M. (2015). When practice takes precedence: conceptions of inquiry and the link to ethical posture. *Educational Action Research, 23*(1), 68-84. https://doi.org/10.1080/09650792.2014.887026

Maier, J., & Richter, T. (2013). Text belief consistency effects in the comprehension of multiple texts with conflicting information. *Cognition and Instruction*, *31*(2), 151-175. https://doi.org/10.1080/07370008.2013.769997

Maier, J., & Richter, T. (2014). Fostering multiple text comprehension: How metacognitive strategies and motivation moderate the text-belief consistency effect. *Metacognition and Learning*, *9*(1), 51-74. https://doi.org/10.1007/s11409-013-9111-x

Mason, L., & Boldrin, A. (2008). Epistemic metacognition in the context of information searching on the Web. In M. S. Khine (Eds) *Knowing, knowledge and beliefs*. (pp. 377-404). Springer.

Mason, L., Boldrin, A., & Ariasi, N. (2010). Searching the Web to learn about a controversial topic: are students epistemically active? *Instructional Science*, *38*(6), 607-633. http://www.jstor.org/stable/23372902

Maxcy, S. J. (2003). Pragmatic threads in mixed methods research in the social sciences: The search for multiple modes of inquiry and the end of the philosophy of formalism. In A. Tashakkori & C. Teddlie (Eds.) *Handbook of mixed methods in social and behavioural research.* (pp51-90). Sage Publications.

McGregor, P. K., Anderson, C. M., Harris, J., Seal, J. R., & Soul, J. M. (1994). Individual differences in songs of fan-tailed warblers Cisticola juncidis in Portugal. *Airo, 5*, 17-21.

Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry, 20*(8), 1045-1053. https://doi.org/10.1177/1077800413513733

Muis, K. R., & Duffy, M. C. (2013). Epistemic climate and epistemic change: Instruction designed to change students' beliefs and learning strategies and improve achievement. *Journal of Educational Psychology, 105*(1), 213-225. https://doi.org/10.1037/a0029690

Muis, K. R., & Franco, G. M. (2009). Epistemic beliefs: Setting the standards for self-regulated learning. *Contemporary Educational Psychology*, *34*(4), 306-318. https://doi.org/10.1016/j.cedpsych.2009.06.005

Newman, S., & Latifi, A. (2021). Vygotsky, education, and teacher education. *Journal of Education for Teaching, 47*(1), 4-17. https://doi.org/10.1080/02607476.2020.1831375

Norton, L. (2007). Pedagogical research in higher education. Ethical issues facing the practitioner-researcher. In A. Campbell & Groundwater-Smith, S. (Eds.) An ethical approach to practitioner research: dealing with issues and dilemmas in action research (pp. 162-171). Routledge.

Nussbaum, E. M., Sinatra, G. M., & Poliquin, A. (2008). Role of epistemic beliefs and scientific argumentation in science learning. *International Journal of Science Education*, 30(15), 1977-1999. https://doi.org/10.1080/09500690701545919

Office for Students (2019). *Preparing for degree study. Analysis of Access to Higher Education Diplomas and integrated foundation year courses.* Office for Students.

Orne, M. T., & Whitehouse, W. G. (2000). Demand characteristics. In A. E. Kazdin (Ed.) *Encyclopaedia of Psychology* (pp. 469–470). American Psychological Association and Oxford University Press.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41*(10), 994-1020. https://doi.org/10.1002/tea.20035

Osborne, J. F., Henderson, J. B., MacPherson, A., Szu, E., Wild, A., & Yao, S. Y. (2016). The development and validation of a learning progression for argumentation in science. *Journal of Research in Science Teaching*, 53(6), 821-846. https://doi.org/10.1002/tea.21316

Osborne, J., Simon, S., Christodoulou, A., Howell-Richardson, C., & Richardson, K. (2013). Learning to argue: A study of four schools and their attempt to develop the use of argumentation as a common instructional practice and its impact on students. *Journal of Research in Science Teaching*, 50(3), 315-347. https://doi.org/10.1002/tea.21073

Perfetti, C. A., Rouet, J.-F., & Britt, A. (1999) Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.) *The construction of mental representations during reading.* (pp. 88-108). Lawrence Erlbaum Associates.

Primor, L., & Katzir, T. (2018). Measuring multiple text integration: A review. *Frontiers in Psychology*, 9, 2294. https://doi.org/10.3389/fpsyg.2018.02294

Quality Assurance Agency. (2019). *Subject Benchmark Statement: Biosciences*. (4th edition). Quality Assurance Agency.

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist*, 52(3), 148-166. https://doi.org/10.1080/00461520.2017.1322968

Ritchie, J., & Spencer, L. (1994). Qualitative data analysis for applied policy research. In A. Bryman & R. G. Burgess (Eds.), *Analyzing qualitative data.* (pp. 173-194). Routledge.

Ritchie, J., & Spencer, L. (2011). Qualitative data analysis for applied policy research. In. A. M. Huberman & M. B. Miles (Eds.), *The Qualitative researcher's companion*. (pp. 305-329). SAGE Publications Inc.

Rosman, T., Mayer, A.-K., Merk, S., & Kerwer, M. (2019). On the benefits of 'doing science': Does integrative writing about scientific controversies foster epistemic beliefs? *Contemporary Educational Psychology, 58*, 85–101. https://doi.org/10.1016/j.cedpsych.2019.02.007

Rouet, J.-F., & Britt, A. (2014). Multimedia learning from multiple documents. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning.* (pp. 813–841). Cambridge University Press.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Sadler, T. D. (2009). Situated learning in science and education: socio-education issues as contexts for practice. *Studies in Science Education*, *45*(1), 1-42. https://doi.org/10.1080/03057260802681839

Sadler, T. D., & Zeidler, D. L. (2005). Patterns of informal reasoning in the context of socioscientific decision-making. *Journal of Research in Science Teaching, 42*(1), 112-138. https://doi.org/10.1002/tea.20042

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, *89*(4), 634-656. https://doi.org/10.1002/sce.20065

Sandoval, W. A., & Bell, P. (2004). Design-based research methods for studying learning in context: Introduction. *Educational Psychologist, 39*(4), 199-201. https://doi.org/10.1207/s15326985ep3904_1

Sandoval, W. A., Greene, J. A., & Bråten, I. (2016). Understanding and promoting thinking about knowledge: Origins, issues, and future directions of research on epistemic cognition. *Review of Research in Education*, *40*(1), 457-496. https://doi.org/10.3102/0091732X1666931

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, *82*(3), 498-504. https://doi.org/10.1037/0022-0663.82.3.498

Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist*, *39*(1), 19-29. https://doi.org/10.1207/s15326985ep3901_3

Schoor, C. (2023). University students' beliefs about science and their relationship with knowledge about science. *European Journal of Psychology of Education*, 1-25. https://doi.org/10.1007/s10212-023-00724-2

Simon, S., Erduran, S., & Osborne, J. (2006). Learning to teach argumentation: Research and development in the science classroom. *International Journal of Science Education, 28* (2-3), 235-260. https://doi.org/10.1080/09500690500336957

Singer, S. R., Nielsen, N. R. & Schweingruber, H. A. (2012). *Discipline-based education research.* National Academies Press.

Smit, J., van Eerde, H. A. A., & Bakker, A. (2013). A conceptualisation of whole-class scaffolding. *British Educational Research Journal, 39*(5), 817-834. https://doi.org/10.1002/berj.3007

Stadtler, M. (2017). The art of reading in a knowledge society: commentary on the special issue on models of multiple text comprehension. *Educational Psychologist*, *52*(3), 225-231. https://doi.org/10.1080/09500690500336957

Stang Lund, E., Bråten, I., Brandmo, C., Brante, E. W., & Strømsø, H. I. (2019). Direct and indirect effects of textual and individual factors on source-content integration when reading about a socio-scientific issue. *Reading and Writing, 32*, 335-356. https://doi.org/10.1007/s11145-018-9868-z

Strømsø, H. I. (2017). Multiple models of multiple-text comprehension: a commentary. *Educational Psychologist*, 52(3), 216-224. https://doi.org/10.1080/00461520.2017.1320557

Strømsø, H. I., Bråten, I., & Britt, M. A. (2011). Do students' beliefs about knowledge and knowing predict their judgement of texts' trustworthiness? *Educational Psychology*, *31*(2), 177-206. https://doi.org/10.1080/01443410.2010.538039

Strømsø, H. I., Bråten, I., Britt, M. A., & Ferguson, L. E. (2013). Spontaneous sourcing among students reading multiple documents. *Cognition and Instruction, 31*(2), 176-203. https://doi.org/10.1080/07370008.2013.769994

Strømsø, H. I., Bråten, I., & Stenseth, T. (2017). The role of students' prior topic beliefs in recall and evaluation of information from texts on socio-scientific issues. *Nordic Psychology*, *69*(3), 127-142. https://doi.org/10.1080/19012276.2016.1198270

Strømsø, H. I., & Kammerer, Y. (2016) Epistemic cognition and reading for understanding in the internet age. In J. A. Greene, W. Sandoval, W., & I. Bråten (Eds.) *Handbook of epistemic cognition*. (pp. 230-246). Routledge.

Swallow, V., Newton, J., & Van Lottum, C. (2003). How to manage and display qualitative data using "Framework" and Microsoft Excel. *Journal of Clinical Nursing, 12*(4), 610–612. https://doi.org/10.1046/j.1365-2702.2003.00728.x

Taylor, L. E., Swerdfeger, A. L., & Eslick, G. D. (2014). Vaccines are not associated with autism: an evidence-based meta-analysis of case-control and cohort studies. *Vaccine, 32*(29), 3623-3629. https://doi.org/10.1016/j.vaccine.2014.04.085

Toulmin, S. (1958) *The uses of argument.* Cambridge University Press.

Trevors, G., Feyzi-Behnagh, R., Azevedo, R., & Bouchet, F. (2016). Self-regulated learning processes vary as a function of epistemic beliefs and contexts: Mixed method evidence from eye tracking and concurrent and retrospective reports. *Learning and Instruction*, *42*, 31-46. https://doi.org/10.1016/j.learninstruc.2015.11.003

Venville, G. J., & Dawson, V. M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. *Journal of Research in Science Teaching*, *47*(8), 952-977. https://doi.org/10.1002/tea.20358

von der Mühlen, S., Richter, T., Schmid, S., & Berthold, K. (2019). How to improve argumentation comprehension in university students: experimental test of a training approach. *Instructional Science*, *47*(2), 215-237. https://doi.org/10.1007/s11251-018-9471-3

Wakefield, A. J., Murch, S. H., Anthony, A., Linnell, J., Casson, D. M., Malik, M., Berelowitz, M., Dhillon, A. P., Thomson, M. A., Harvey, P., Valentine, A., Davies, S. E., & Walker-Smith, J. A. (1998). RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, *351*(9103), 637-641. https://doi.org/10.1016/s0140-6736(97)11096-0

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*, *46*(4), 1060-1106. https://doi.org/10.3102/0002831209333183

Wiley, J., & Voss, J. F. (1996). The effects of 'playing historian' on learning in history. *Applied Cognitive Psychology*, *10*(7), 63-72. https://doi.org/10.1002/(SICI)1099-0720(199611)10:7%3C63::AID-ACP438%3E3.0.CO;2-5

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial

evidence. *Journal of Educational Psychology*, *83*(1), 73. https://doi.org/10.1037/0022-0663.83.1.73

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89-100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Appendices

Appendix A: The Justification for Knowing Questionnaire

Student ID number		Anonymised code (to be completed by researcher)	

The following questions are about the natural sciences. There are no correct or incorrect answers, because I am interested in your personal beliefs.

Use the scale below to answer the questions. If you agree completely with a statement, then put a cross in the box under number 10 for that statement; if you disagree completely, put a cross in the box under number 1 for that statement. If you are more or less in agreement with a statement, put a cross in the box between 1 and 10 that best expresses your opinion. Put one cross only for each statement.

		Disagree completely									Agree completely
		1	2	3	4	5	6	7	8	9	10
1	To decide whether something I read about the natural sciences is correct, I have to check whether it is in accordance with other things I have read or heard about natural science										
2	When I read something about the natural sciences that is based on scientific investigations, then I believe that it is correct										
3	Knowledge about the natural sciences is only personal opinion – there are no facts										
4	Just one source is never enough to decide what is right in the natural sciences										
5	I can never be sure about a claim in the natural sciences until I have checked it with at least one other source										
6	Everyone can have different opinions about the natural sciences because no completely correct answers exist										
7	If a science teacher says something is correct, then I believe it										
8	Things that are written in natural sciences textbooks are correct										
9	To be able to trust knowledge claims in the natural sciences, I have to check various knowledge sources										
10	I believe in claims that are based on scientific research										
11	To detect incorrect claims in texts about the natural sciences, it is important to check several information sources										
12	I believe that everything I learn in science class is correct										
13	What is a fact in the natural sciences depends on one's personal view										
14	If a scientist says that something is a fact, then I believe it										

Appendix B: Study 1 Interview Protocol

Thank you for agreeing to take part in this interview. During the next 30 minutes, I'd like you to tell me about your experiences of completing the Making an argument assignment.

1. Tell me about the issue you chose.
    a. Why did you choose this issue?
    b. What did you already know about this issue?
    c. Did you have strong views about this issue at the start? Did your views change?
    d. Were you interested in knowing what experts thought about this issue?
    e. Whose views were more important? Yours or the experts? What do you think now?

2. Before you started the assignment, there was some preparatory work.
    a. Can you tell me what you found most helpful and what you found least helpful?
    b. What do you think I should use to help next year's students prepare? Why is that?

3. I'd now like to know about your experiences of using multiple online sources to construct an argument.
    a. Can you tell me about how you looked for sources? What were you looking for?
    b. How did you decide which of these sources you'd use in your assignment?
    c. Did the sources agree or disagree with what you already knew / your own views?
    d. Did you look for sources which contradicted each other?
    e. If sources were contradictory, how did you deal with this?
    f. If you're asked to do a similar piece of work in the future, will you look at lots of sources or just a few? Why is that?
    g. Did you check whether the sources were written by experts? If so, how? If not, why not?

Thanks again for agreeing to take part. It's been really interesting to talk to you.

Appendix C: Study 2 Interview Protocol

Thank you for agreeing to take part in this interview. Over the next 20-30 minutes, I'd like you to tell me about your experiences of activities we did during the workshops at the start of the spring semester to help you complete the 'Making an argument' assignment.

I'm going to start by reminding you about the first workshop. During that workshop, we looked at the controversy surrounding the MMR vaccine and whether it causes autism to introduce the idea that there can be different perspectives about socio-scientific issues.

- First, we talked about how the views of Andrew Wakefield differed from the rest of his research team.
- Then there were a series of activities. For one of these, you had to read a small section from the Lewis and Speers article which said that the MMR and autism controversy was one of the first times that parents were used as 'experts' within the media, and so we talked about whether anecdotal evidence from individuals should be regarded in the same way as evidence from scientific research.
- Another of the activities looked at a meta-analysis which combined the results from many scientific studies about whether there is a link between the MMR vaccine and autism.
- There was also an activity where you had to decide whether websites were 'trustworthy'.

1. Did these activities help you to understand that there can be different perspectives about an issue and if so, how? If not, what could I do differently next time to help students understand this?

2. Did the activities in this workshop help you when you were choosing your own issue and if so, how?
    a. Did you know much the issue you chose before you started reading about it?
    b. Did you hold strong personal views? Did these views change?
    c. Were you aware of any other perspectives about this issue?

During the next couple of workshops, we looked at how to construct an argument. The first workshop dealt with the claim, evidence, and warrant. The second workshop was about how to refine your argument using qualifiers, rebuttals, and backings.

3. Were you familiar with how to construct arguments before these workshops?
    a. What did you learn from these workshops?
    b. Was it helpful to look at edited extracts of past students' work during these workshops? If so, why was that?
    c. Is there anything I could differently next time to help students learn how to construct an argument?

During the final couple of workshops, we thought about how to find and read journal articles. You then had to select and annotate three articles and we then had a tutorial where we talked about how you were going to use information from articles to construct your assignment.

4. Were you interested in whether the people who'd written the articles were 'experts'? If so, how did you check whether they were experts?

5. What did you learn during the workshops and the tutorial about how to use multiple articles to construct your argument? Is there anything I should do differently next year?

6. Finally, is there anything else that you'd like to tell me about your experiences of any of the workshops or the individual tutorials?

Appendix D: Information Sheet for Participants

The experience of constructing a balanced argument from multiple online documents

Who is conducting the study?
My name is Caroline Anderson. I am a postgraduate research student on the Professional Doctorate in Education programme in the School of Education at The University of Nottingham.

What is the project about?
I am interested in your experiences of constructing a balanced argument from multiple online documents.

What are you being asked to do?
The planned research involves collecting data about your experiences of the educational activities which are a normal part of the Studying Science at University module in the spring semester.

You are being asked if you would be willing to participate in the data collection activities. There are five data collection activities which are organised into two stages. You can choose whether you like to participate in both stages of data collection, or in stage one only. You do not have to participate in any stages of the data collection if you do not wish to do so.

Stage One
Stage one features two data collection activities. If you agree to take part in stage one, you are giving consent to participating in both of these activities:

1. You will be asked to fill in a short questionnaire about your beliefs about how you know something. You will be asked to fill in this questionnaire twice; once at the start of the spring semester and again just after the Easter break.

2. The marks you are awarded for the 'Making an argument' assignment will be analysed.

Stage Two
Stage two features three data collection activities. If you agree to take part in stage two, you are giving consent to participating in all three of these activities:

1. During the spring semester, you will have an individual tutorial about your choice of sources for the 'Making an argument' assignment. I will be recording all of these conversations so that you can listen back when you are writing your 'Reflective commentary' assignment. I will be analysing these conversations.

2. The responses you give in your 'Reflective commentary' assignment will be analysed.

3. A small number of students will be invited to participate in individual 30-minute interviews to talk about their experiences of constructing a balanced argument from multiple online documents.

What are the other things you need to know?
You can ask me any questions, at any time, about my research.

If you consent to me doing so, I will keep a copy of your 'Reflective Commentary' assignment so that I can analyse what you have written.

All data, which includes signed consent forms, questionnaires, copies of assignments, recordings, transcriptions and analyses of individual tutorials and interviews, will be destroyed after seven years.

All data, with the exception of consent forms, will be anonymised for storage. If I write anything about you, I will give you a different name so that no-one will know who you are.

Do you have to take part?
All students are expected to take part in the educational activities which are a normal part of the Studying Science at University module in the spring semester.

You do not have to take part in any data collection activities. You can stop at any point if you no longer wish to take part in the data collection activities.

If you do not wish to take part, or if you wish to withdraw part-way through the project, you do not have to give a reason and there will not be any consequences.

Contact details

Researcher: Caroline Anderson, EdD candidate
caroline.anderson1@nottingham.ac.uk

Supervisor:
Professor Shaaron Ainsworth, School of Education
shaaron.ainsworth@nottingham.ac.uk

School of Education Research Ethics Coordinator:
educationresearchethics@nottingham.ac.uk

Appendix E: Participant Consent Form

The experience of constructing a balanced argument from multiple online documents

- I have read the Participant Information Sheet and the nature and purpose of the research project has been explained to me.
- I understand the purpose of the research project and my involvement in it.
- I understand that I am able to choose whether to be involved in both stages of data collection or stage one only.

I consent to take part in:

Both stage one and stage two of data collection	
Stage one only of data collection	

- I understand that I may withdraw from the research project at any stage and that this will not affect my status now or in the future.
- I understand that while information gained during the study may be published, I will not be identified and my personal results will remain confidential.
- I understand that I will be audio recorded during my individual tutorial and interview.
- I understand that data will be stored both as hard copies and electronically. Hard copy files will be stored in a locked filing cabinet in a locked office. Electronic data files will be password protected and stored on secure server with access limited to the researcher and supervisors. Data will be deleted 7 years after any publication based upon it.
- All participant names will be changed, and any other identifiers, such as student identification numbers, will be replaced with participant numbers.
- I understand that I may contact the researcher or supervisor if I require further information about the research, and that I may contact the Research Ethics Coordinator of the School of Education, University of Nottingham, if I wish to make a complaint relating to my involvement in the research.

Signed …… (research participant)

Print name …… Date ……

Contact details

Researcher: Caroline Anderson, EdD candidate
caroline.anderson1@nottingham.ac.uk

Supervisor: Professor Shaaron Ainsworth
shaaron.ainsworth@nottingham.ac.uk

School of Education Research Ethics Coordinator:
educationresearchethics@nottingham.ac.uk

Appendix F: GDPR Full Privacy Notice

Full Privacy Notice for Research Participants

How the University of Nottingham processes your personal data
The University of Nottingham, University Park, Nottingham, NG7 2RD (0115 951 5151), is committed to protecting your personal data and informing you of your rights in relation to that data.

The University of Nottingham is registered as a Data Controller under the Data Protection act 1998 (registration No. **Z5654762 –** https://ico.org.uk/ESDWebPages/Entry/Z5654762).

One of our responsibilities as a data controller is to be transparent in our processing of your personal data and to tell you about the different ways in which we collect and use your personal data. The University will process your personal data in accordance with the General Data Protection Regulation (GDPR) and the Data Protection Act 2018 and this privacy notice is issued in accordance with the GDPR Articles 13 and 14.

We may update our Privacy Notices at any time. The current version of all of our Privacy Notices can be found at https://www.nottingham.ac.uk/utilities/privacy.aspx and we encourage you to check back regularly to review any changes.

The Data Protection Officer
The University has appointed a Data Protection Officer. Their postal address is:
Data Protection Officer,
Legal Services
A5, Trent Building,
University of Nottingham,
University Park,
Nottingham
NG7 2RD
They can be emailed at dpo@nottingham.ac.uk.

Your personal data and its processing
We define personal data as information relating to a living, identifiable individual. It can also include "special categories of data", which is information about your racial or ethnic origin, religious or other beliefs, and physical or mental health, the processing of which is subject to strict requirements. Similarly information about criminal convictions and offences is also subject to strict requirements. “Processing” means any operation which we carry out using your personal data e.g. obtaining, storing, transferring and deleting.

We only process data for specified purposes and if it is justified in accordance with data protection law. Detail of each processing purpose and its legal basis is given in each privacy notice listed below, please select the one most relevant to your relationship to the University.

Why we collect your personal data
We collect personal data under the terms of the University's Royal Charter in our capacity as a teaching and research body to advance education and learning. Specific purposes for data collection on this occasion are to design, implement and evaluate learning and teaching activities that enable students with differing initial epistemic beliefs to construct an argument about a socio-scientific issue using multiple online documents.

Legal basis for processing your personal data under GDPR
The legal basis for processing your personal data on this occasion is Article 6(1e): processing is necessary for the performance of a task carried out in the public interest.

How long we keep your data
The University may store your data for up to 25 years and for a period of no less than 7 years after the research project finishes. The researchers who gathered or processed the data may also store the data indefinitely and reuse it in future research. Measures to safeguard your stored data include pseudonymisation procedure, and anonymisation of data. Data will be treated confidentially, and only the researcher and the researcher's supervisor will have access to the raw data. All information collected while carrying out the study will be stored at a location which is password protected and strictly confidential. Any resulting digital or textual data will also be stored electronically on a password protected hard drive for a period of seven years after any research is published before being destroyed. The proposed research project has been approved by the University of Nottingham Ethics committee in compliance with the University's Code of Research Conduct and Research Ethics and follows the revised ethical protocols outlined by the British Educational Research Association (2018). This means the School's name, and students' identity will be anonymised from the outset before digital storage and all materials relating to the School and students will be confidential and only available to my supervisor and me.

Who we share your data with
Extracts of your data may be disclosed in published works that are posted online for use by the scientific community. Your data may also be stored indefinitely on external data repositories (e.g., the UK Data Archive) and be further processed for archiving purposes in the public interest, or for historical, scientific or statistical purposes. It may also move with the researcher who collected your data to another institution in the future. Following analysis, data will be used for my EdD thesis and within research publications such as journals, book chapters and conference proceedings. Additionally, data may be shared within the researcher's School and university. This is because it is hoped that the findings of this research will enable

others to design, implement and evaluate learning and teaching activities which enable students with differing initial epistemic beliefs to construct balanced arguments about controversial issues using multiple online documents.

Your rights as a data subject
You have the following rights in relation to your personal data processed by us:

Right to be informed
The University will ensure you have sufficient information to ensure that you're happy about how and why we're handling your personal data, and that you know how to enforce your rights.

The University provides information in the form of privacy notices. Our Privacy Notices pages can be found at https://www.nottingham.ac.uk/utilities/privacy/privacy.aspx.

Right of access / right to data portability
You have a right to see all the information the University holds about you. Where data is held electronically in a structured form, such as in a database, you have a right to receive that data in a common electronic format that allows you to supply that data to a third party - this is called "data portability".

To make a request for your own information please see the link here: https://www.nottingham.ac.uk/governance/records-and-information-management/data-protection/data-protection.aspx

To receive your information in a portable form, send an email your request to data-protection@nottingham.ac.uk

Right of rectification
If we're holding data about you that is incorrect, you have the right to have it corrected.

Please email any related request to data-protection@nottingham.ac.uk.

Right to erasure
You can ask that we delete your data and where this is appropriate we will take reasonable steps to do so.
Please email any related request to data-protection@nottingham.ac.uk.

Right to restrict processing
If you think there's a problem with the accuracy of the data we hold about you, or we're using data about you unlawfully, you can request that any current processing is suspended until a resolution is agreed.

Please email any related request to data-protection@nottingham.ac.uk.

Right to object
You have a right to opt out of direct marketing.
You have a right to object to how we use your data if we do so on the basis of "legitimate interests" or "in the performance of a task in the public interest" or "exercise of official authority" (a privacy notice will clearly state to you if this is the case). Unless we can show a compelling case why our use of data is justified, we have to stop using your data in the way that you've objected to.

For direct marketing, there will be an opt-out provided at the point of receipt. To object to how we use your data, email your request to data-protection@nottingham.ac.uk.

Rights related to automated decision making including profiling
We may use a computer program, system or neural network to make decisions about you (for example, everyone that is on a particular course gets sent a particular letter) or to profile you. You have the right to ask for a human being to intervene on your behalf or to check a decision.

Please email any related request to data-protection@nottingham.ac.uk.

Withdrawing consent
If we are relying on your consent to process your data, you may withdraw your consent at any time.

Exercising your rights, queries and complaints
For more information on your rights, if you wish to exercise any right, for any queries you may have or if you wish to make a complaint, please contact our Data Protection Officer.

Complaint to the Information Commissioner
You have a right to complain to the Information Commissioner's Office (ICO) about the way in which we process your personal data. You can make a complaint on the ICO's website.

Privacy notices
Please consult the privacy notice that best fits your relationship with the University.

Appendix G: Research Participant Privacy Notice

Privacy information for Research Participants
For information about the University's obligations with respect to your data, who you can get in touch with and your rights as a data subject, please visit: https://www.nottingham.ac.uk/utilities/privacy.aspx.

Why we collect your personal data
We collect personal data under the terms of the University's Royal Charter in our capacity as a teaching and research body to advance education and learning. Specific purposes for data collection on this occasion are design, implement and evaluate learning and teaching activities that enable students with differing initial epistemic beliefs to construct an argument about a socio-scientific issue using multiple online documents.

Legal basis for processing your personal data under GDPR
The legal basis for processing your personal data on this occasion is Article 6(1e): processing is necessary for the performance of a task carried out in the public interest.

How long we keep your data
The University may store your data for up to 25 years and for a period of no less than 7 years after the research project finishes. The researchers who gathered or processed the data may also store the data indefinitely and reuse it in future research. Measures to safeguard your stored data include pseudonymisation procedure, and anonymisation of data. Data will be treated confidentially, and only the researcher and the researcher's supervisor will have access to the raw data. All information collected while carrying out the study will be stored at a location which is password protected and strictly confidential. Any resulting digital or textual data will also be stored electronically on a password protected hard drive for a period of seven years after any research is published before being destroyed. The proposed research project has been approved by the University of Nottingham Ethics committee in compliance with the University's Code of Research Conduct and Research Ethics and follows the revised ethical protocols outlined by the British Educational Research Association (2018). This means the School's name, and students' identity will be anonymised from the outset before digital storage and all materials relating to the School and students will be confidential and only available to my supervisor and me.

Who we share your data with
Extracts of your data may be disclosed in published works that are posted online for use by the scientific community. Your data may also be stored indefinitely on external data repositories (e.g., the UK Data Archive) and be further processed for archiving purposes in the public interest, or for historical, scientific or statistical purposes. It may also move with the researcher who collected your data to another institution in the future. Following analysis, data will be used for my EdD thesis and within research publications such as journals, book chapters and conference

proceedings. Additionally, data may be shared within the researcher's School and university. This is because it is hoped that the findings of this research will enable others to design, implement and evaluate learning and teaching activities which enable students with differing initial epistemic beliefs to construct balanced arguments about controversial issues using multiple online documents.

Appendix H: Dendrograms of the Study 1 Participants

Figure 10-1 *Pre-Intervention Study 1 Clusters*

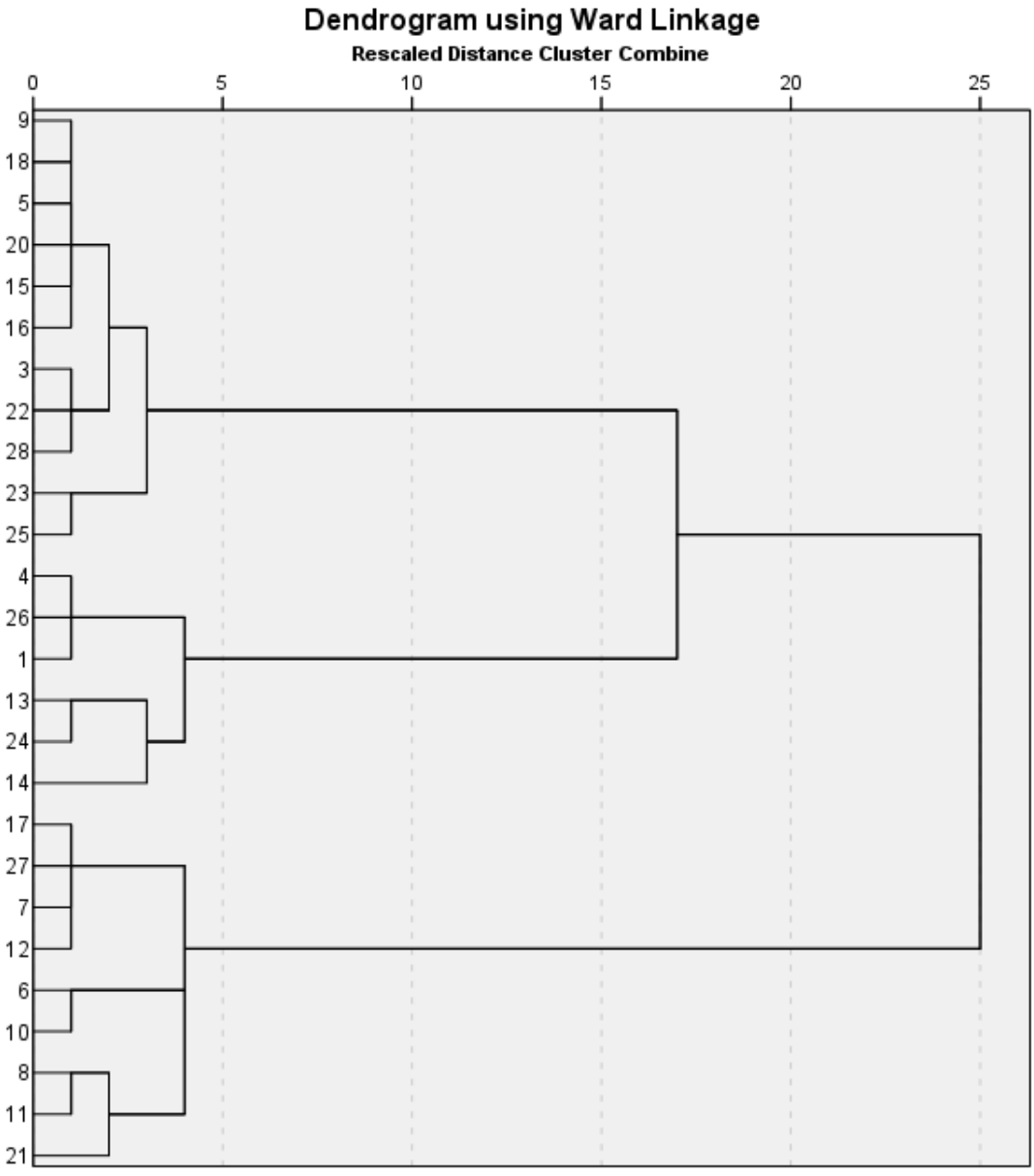
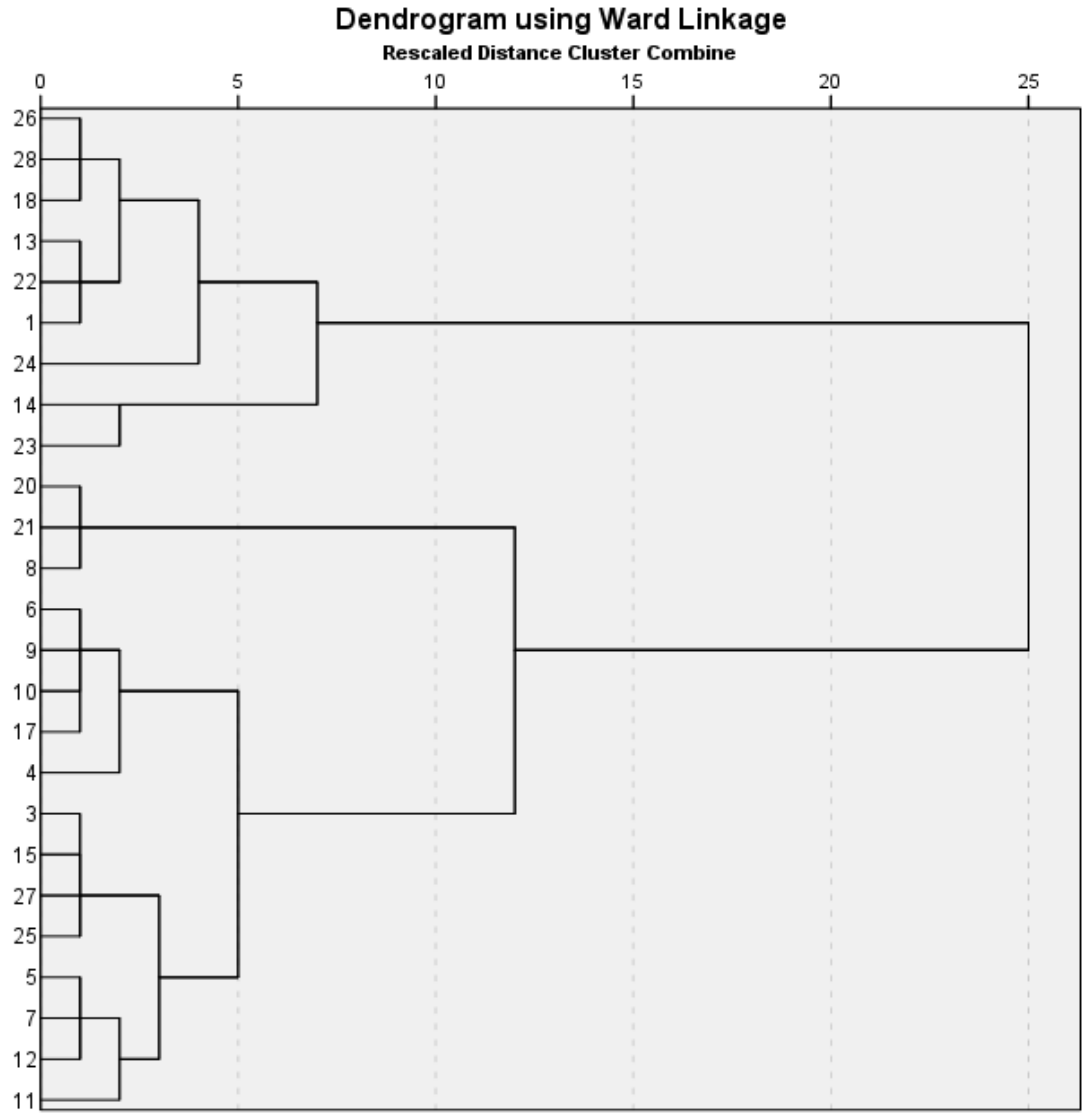


Figure 10-2 *Post-Intervention Study 1 Clusters*



Appendix I: Dendrograms of the Study 2 Participants

Figure 10-3 *Pre-Intervention Study 2 Clusters*

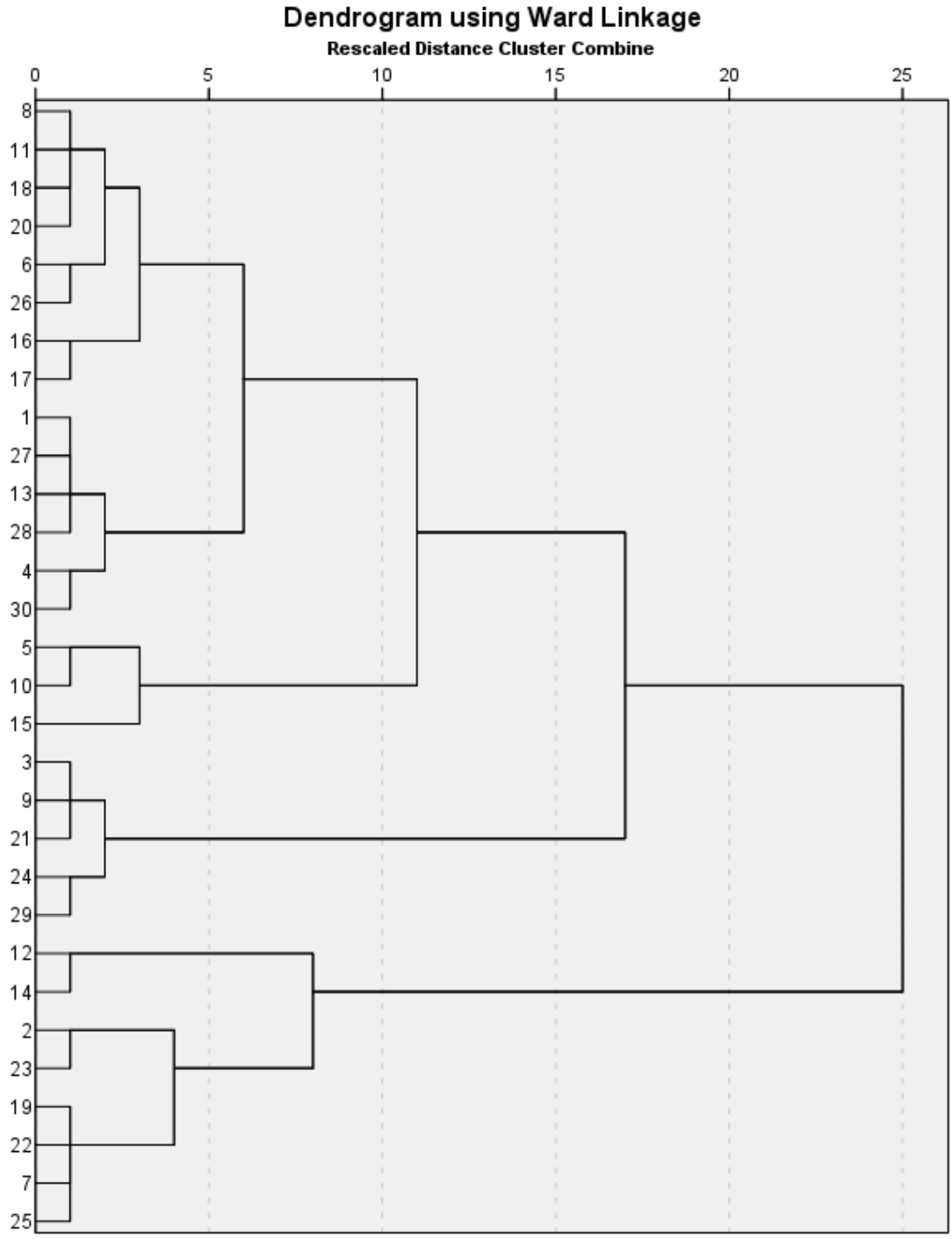


Figure 10.4 *Post-Intervention Study 2 Clusters*