Developing Year 10 History students' written causal arguments through a focus on their conceptual understanding and self-regulation

Alex Hawksworth-Brookes

A Research & Development Project

Submitted for the MSc Learning & Teaching 2020

DEPOSIT AND CONSULTATION OF THESIS

One copy of your dissertation will be deposited in the Department of Education Library via the WebLearn site where it is intended to be available for consultation by all Library users. In order to facilitate this, the following form should be completed which will be inserted in the library copy your dissertation.

Note that some graphs/tables may be removed in order to comply with copyright restrictions.

Surname	Hawksworth-Brookes
First Name	Alex
Faculty Board	Education
Title of Dissertation	Developing Year 10 History students' written causal arguments through a focus on their conceptual understanding and self-regulation

Declaration by the candidate as author of the dissertation

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.

2. I understand that the Department requires that I shall deposit a copy of my dissertation in the Department of Education Library via the WebLearn site where it shall be available for consultation, and that reproductions of it may be made for other Libraries so that it can be available to those who to consult it elsewhere. I understand that the Library, before allowing my dissertation to be consulted either in the original or in reproduced form, will require each person wishing to consult it to sign a declaration that he or she recognises that the copyright of this thesis belongs to me. I permit limited copying of my dissertation by individuals (no more than 5% or one chapter) for personal research use. No quotation from it and no information derived from it may be published without my prior written consent and I undertake to supply a current address to the Library so this consent can be sought.

3. I agree that my dissertation shall be available for consultation in accordance with paragraph 2 above.

Abstract

This project evolved from a desire to improve students' written causal arguments, whilst also improving their metacognition and self-regulation. Baseline data, collected from student essays, a talk-through protocol, and a survey, confirmed initial hypotheses that students had a limited understanding of causation, found constructing causal arguments in History challenging, and possessed limited self-regulatory skills and techniques. An action research model was followed, with three teaching interventions interspersed with analysis of students' work. The first phase used 'link diagrams' to cultivate an understanding of causation as 'web-like'; the second used 'circle diagrams' to focus students' attention on the need to reach decisions about the relative importance of different causes and provide them with a planning tool to promote self-regulation; and the third targeted the use of causation-specific language. Analysis of students' work following each intervention and comparison with both the baseline data and the views of non-intervention class (used as a form of control group) indicated that the interventions had a positive impact, particularly on student understanding of causation. This was confirmed by a final round of data collection and analysis, using the same three sources as for the baseline. Whilst promising progress was made by most students in terms of their causal understanding and written arguments, and by some in terms of their self-regulation, further refinements to the three teaching interventions are needed in future years.

Acknowledgments

Sincere and significant thanks go, first and foremost, to my supervisor, Dr Katharine Burn, for the many hours spent discussing all things cause and consequence and self-regulation, both in person and virtually, and for guiding me through a project that at times felt overwhelming. I am also grateful to Richard Stamper for his help with the statistical analysis of my survey data, and to Hugh Richards for sharing with me an early draft of an article on self-regulatory planning techniques, which provided significant inspiration for my second teaching intervention. I would also like to acknowledge the support provided to me by my school, both in terms of their contribution to my funding, and in their support for my research. I am grateful to the various friends, family, and colleagues who have proofed various drafts and provided suggestions for improvements. Finally, thank you to my students, without whom this research would not have been possible.

Table of Contents

Abstract page 2

Acknowledgements page 3

Table of Contents page 4

List of Figures page 6

List of Tables page 6

Rationale page 7

Literature Review page 9

1. Causation page 9

a. What does effective causal arguing in History look like? page 10

b. What do students find challenging when constructing causal arguments? page 13

c. What does research suggest about how to help students improve their understanding of causation? page 16

2. Self-regulation page 21

a. What is self-regulation and why is it important? page 22

b. What are the barriers to effective self-regulation? page 23

c. What does research suggest about how to help improve student self-regulation? page 25

Implications and Research Questions page 28

Methodology: Overview page 30

Baseline: initial data collection methodology and ethics page 32

a. Student essays page 32

b. Think-out-loud exercise page 33

c. Student survey page 34

Baseline data findings page 36

a. Student essays page 36

b. Think-out-loud exercise page 37

c. Student survey page 38

Implications for teaching interventions page 42

1a: Task Design: Link Diagrams page 43

1b: Intervention 1: analysis, findings, and reflections page 44

Coding methodology page 44

"Web-like" explanations page 45

Causal arguments page 46

Other findings and reflections page 47

2a: Task Design: Circle Diagrams page 47

2b: Intervention 2: analysis, findings, and reflections page 49

3a: Task Design: Vocabulary-targeting Exercises page 51

3b: Intervention 3: analysis, findings, and reflections page 52

E: Evaluation of project as a whole page 55

Methodology page 55

a. Student work page 55

b. Think-out-loud exercises page 56

c. Student surveys page 56

Findings page 57

a. Student work page 57

b. Think-out-loud exercises page 59

c. Student surveys page 60

Final reflections and next steps page 66

Reference List page 69

Appendix A: Initial Student Survey page 76

Appendix B: CUREC documentation page 80

Appendix C: Link Diagram template: Hitler's appointment as Chancellor page 83

Appendix D: Example of coding of work from Intervention 1 page 84

Appendix E: Illustrative examples of coding of pupil work page 85

Appendix F: Link Diagram Template: Development of the Cold War page 86

Appendix G: Exemplar Circle Diagram from Intervention 2 page 87

Appendix H: Example of coding of work from Intervention 2 page 88

Appendix I: An intervention class student's Circle Diagram completed during Intervention 2, with self-regulatory ticking of content included in her essay page 89

Appendix J: Resources provided to students in Intervention 3 page 90

Appendix K: Example of coding of work from end of project analysis page 95

Appendix L: Endpoint student survey page 96

Appendix M: Two-tailed t-test data page 101

List of Figures

Figure 1: summaries of Lee and Shelmilt's (2009) and Scott's (1990) progression models: page 13

Figure 2: EEF recommendations for developing metacognition and self-regulated learning: page 26

Figure 3: EEF metacognitive self-regulation strategies: page 27

Figure 4: Project overview: page 31

List of Tables

Table 1: Baseline survey responses (Q1-8 & Q10): page 39

Table 2: Baseline survey responses (Q9, Q11 & Q12): page 40

Table 3: Student use of causal vocabulary: page 53

Table 4: Endpoint survey responses (Q1-6, 8-11 & 13-14): page 61

Table 5: Endpoint survey responses (Q7): page 62

Table 6: Endpoint survey responses (Q12): page 62

Table 7: Endpoint survey responses (Q15-16): page 62

Table 8: Endpoint survey responses (Q17-21): page 63

Rationale

The inspiration for this study originated from conversations with colleagues in my school's History Department. We regularly bemoaned students' reluctance to think for themselves and how quickly they asked, "What do *you* think the right argument is?" which also revealed a worrying misconception about the nature of History as a discipline (Lee and Shemilt, 2009). Consulting other departments revealed that this manifested as a school-wide phenomenon. For many teachers, the solution had been to provide students with 'model' answers and writing frames, despite our reservations about the potential long-term negative impact on students' self-sufficiency. It was subsequently reassuring to read of similar frustrations among other teachers, as reported within the History education community, and within a series of CPD sessions early in the 2019/20 academic year, where members of the SLT confirmed that student metacognition was to become an area of focus. Therefore, some form of research into this topic seemed pertinent, from both a departmental and school-wide perspective.

Furthermore, in History, my colleagues and I felt that students' written work could be improved. Although my school performs highly in public exams (100% 7 to 9 in GCSE History in 2019), we agreed that students' written work often lacked finesse and complexity, particularly in relation to constructing a well-substantiated argument, featuring high-quality links and comparisons between factors. My Head of Department expressed the belief that this often resulted in students not reaching the highest bands at GCSE and was therefore an area where improvements could be made. This helped to refine my focus to be on students' written arguments and to identify a specific cohort of students: those currently studying GCSE History. This decision was also because, as a department, we knew that students who continued with History at A Level often found the step up a challenging one: a concern shared by other A-level teachers (Carroll, 2016a). Improving students' cognitive independence and ability to argue at GCSE would hopefully ease this transition. Furthermore, GCSE seemed an appropriate stage for this project, as students were required to construct complex

arguments far more regularly than at Key Stage 3. This would provide ample opportunity to collect and analyse data and implement a series of interventions.

Finally, I chose to focus exclusively on causation. My school's specification (Edexcel IGCSE) places a significant emphasis on cause and consequence. Across the two papers, three compulsory shorter questions are dedicated to it, whilst three out of four extended essay questions regularly focus on causation, meaning that up to 60% of a student's total GCSE mark can involve cause and consequence. Additionally, focusing on a specific concept would allow me to develop more-targeted interventions with which to support my students.

I therefore devised a study centred on two key objectives:

1. Improving students' understanding of causation and therefore their ability to construct complex and convincing written causal arguments.
2. Improving students' independence as historical thinkers through a focus on their self-regulation.

As a result, I established a series of questions to guide my literature review:

1. Causation
    a) What does effective causal arguing in History look like?
    b) What do students find challenging when constructing causal arguments?
    c) What does research suggest about how to help students improve their understanding of causation?
2. Metacognition
    a) What is self-regulation and why is it important?
    b) What are the barriers to effective self-regulation?
    c) What does research suggest about how to help improve student self-regulation?

I hoped that a review framed by these questions would help me to give my students the tools and understanding to argue more effectively about causation and thereby become more self-sufficient learners.

Literature Review

1. Causation

Any discussion of causation would be incomplete without reference to influential historian Carr's (1986) ubiquitous statement that "the study of history is a study of causes" (p.81). Although recent shifts in historiographical trends, particularly in terms of the increased profile of concepts such as diversity, have challenged this, cause and consequence are still central to much of school-level History (Woodcock, 2010; Watts, 2016). Furthermore, it is a strongly held view that a secure understanding of the key historical concepts themselves is a vital part in developing students' historical understanding as a whole, and that the substantive and conceptual cannot be truly separated, indicating that dedicating lesson time to them is an important part of History teaching (Haydn et al., 2015; Chapman, 2017). Much of the following research, particularly regarding teaching interventions, comes from small-scale practitioner studies, reflecting the nature of research in the history community, with a limited amount of large-scale empirical research (Burn et al., 2019). However, there are some significant exceptions, such as the work of Lee and Shemilt whose large-scale empirical work – evaluating the Schools Council History Project (Shemilt 1980) and exploring children's concepts of history (Ashby et al., 1997) – has contributed to much of our understanding of students' knowledge and misconceptions about causation. Therefore, whilst much of what follows is valuable, coming from practising teachers with high engagement with their subject and the historical community, it is important to bear in mind that many of the interventions are small-scale and context-specific.

a. What does effective causal arguing in History look like?

Relationships are central to the concept of cause and consequence. As Shemilt (1986), in a large-scale review of the Schools History Project, states, causation is 'a concept defining the relation obtaining between events' (p.6). Understanding the **interplay** of different events, individuals' actions, contexts, beliefs, etc. in bringing about change is central to students' progression. As my emphasis demonstrates, it is the relationship *between* things that is important (Evans, 1997; Historical Association, 2019).

Effectively discussing causation requires various forms of categorisation, which vary in both form and cognitive complexity. A relatively straightforward, albeit still important, method is descriptive; i.e. grouping causal relationships according to themes such as "social", "political", and "economic"; similarly, temporal categorisation ("long-term", "short-term") is another descriptive classification that can be used to organise information (Chapman, 2003; Wineburg, 2007; Morton and Seixas, 2012). However, such forms of labelling do not necessarily precipitate effective causal arguments, as they merely *describe* rather than explain the *impact* of events (Chapman, 2017; Historical Association, 2019).

It is self-evident, therefore, that more effective causal argument requires some form of explanatory classification as well. Simply stating that X led to Y demonstrates superficial understanding at best. As discussed in the following section, moving to a more complex form of labelling can be challenging for students. One such approach is Coffin's (2006) schema of enabling and determining relationships. Verbs and adverbs can be powerful tools in explaining enabling relationships, through which the likelihood of future events or states of affairs is affected (e.g. "X *accelerated* discontent"); terms such as "trigger", "catalyst", and "precondition" are equally useful (Lomas, 1993; Chapman, 2003; Woodcock, 2005). Such terminology also elucidates understanding of determining relationships, these being explanations of a cause's role in shaping an event's time, place, nature, etc. (Coffin, 2006; Chapman 2017). Importantly, these types of causal relationship are not parameters that exist independently; complex causal arguments will blend them, demonstrating

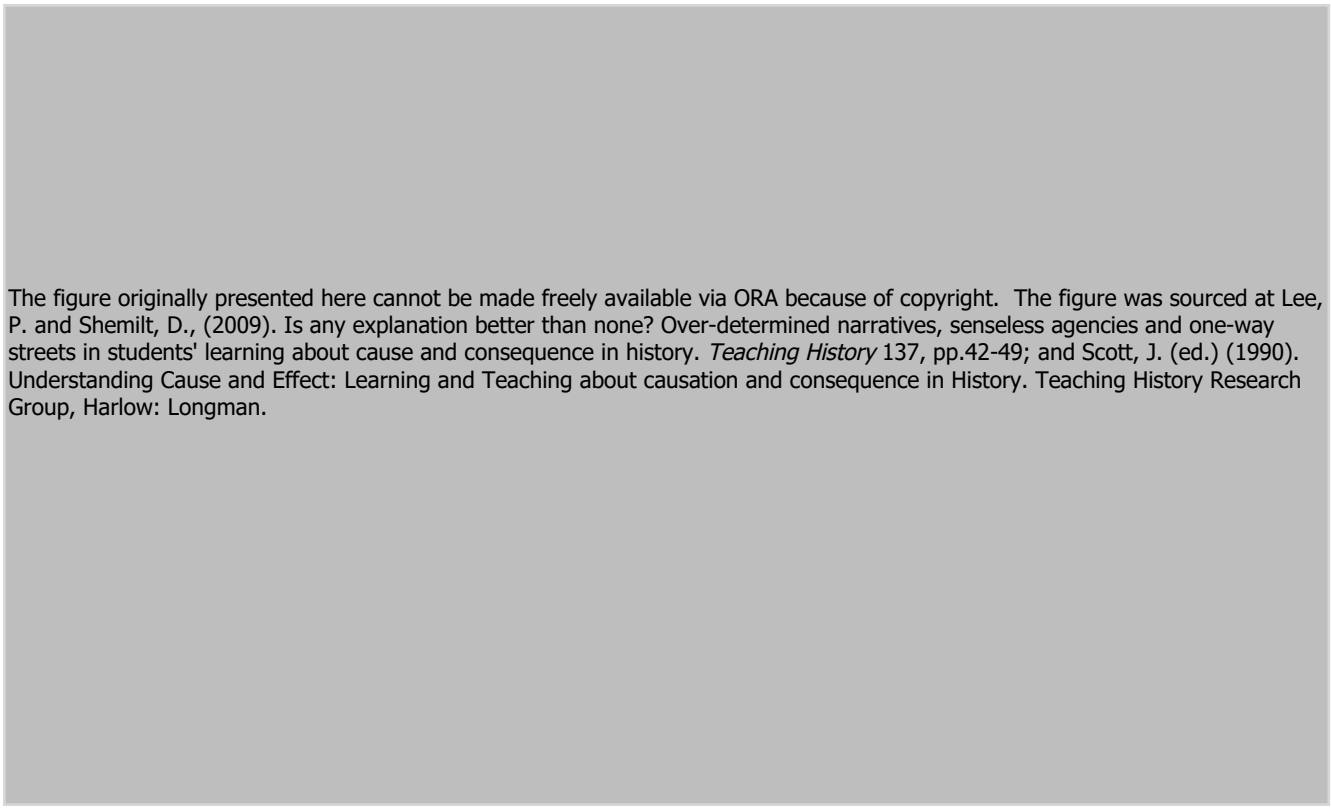
sophisticated understanding in the process. Further complexity comes from the interplay between individual and collective action. Empathy is important in helping students understand aspects of human behaviour such as collective mentalities, agency, and individual action (McCullagh, 2004; Woodcock, 2010; Lee and Shemilt, 2011; Worth, 2011).

Such explanatory forms of classification facilitate making judgements about relative importance or, in other words, evaluating causal relationships. The high levels of challenge associated with this process are reflected in exam board mark schemes. For example, the generic mark scheme provided by my school's GCSE board for their extended essay question lists the following as a requirement for students to reach the highest mark band: "Criteria for the required judgement are **justified** and applied in the process of reaching the overall judgement" (Pearson Edexcel, 2016 – emphasis added). Indeed, the wording of Edexcel's IGCSE essay questions places the idea of relative importance at the heart of the question, as an example from the 2019 summer paper demonstrates: "'Party reorganisation was the **main reason** why the Nazi Party survived in the years 1924–28.' **How far** do you agree? Explain your answer." (Pearson Edexcel, 2019 – emphasis added). The role of such hierarchical classification is not confined to exam boards; the noted historian Evans (1997) writes of a wide range of types of cause, including "necessary", "sufficient", "absolute", and "relative"; concepts which can allow historians to connect and rank factors (pp.157-158). Indeed, such thinking is what lies behind Carr's (1986) "true historian['s] … final analysis" (p.84).

The above features form a natural progression, from description to explanation to evaluation, which mirrors the hierarchy of skills seen in assessment materials. A range of research-based models have also been suggested to help categorise causal explanations produced by students. Lee and Shemilt (2009) proposed a six-level model, based on a wide-ranging analysis of the qualities evident in student work (see Figure 1, below); in the lower stages, students' thinking tends to be constrained by misconceptions about the nature of human agency and what they perceive as the deterministic nature of history (pp.43-45). At Level 4 a "re-conceptualis[ation]" occurs, whereby students become capable of "possibility thinking" – i.e. the ability to acknowledge other historical possibilities (pp.45-

48). For Lee and Shemilt, this level of their model is one that "a minority of 14–18-year-old students make the breakthrough into", with Level 5 and the conjectural Level 6 being increasingly abstract and distinct from the models of progression outlined by exam boards. (pp.47-48). Scott's (1990) model bear similarities with Lee's and Shemilt's (2009) whilst acting as a more conventional progression model; again, the lower levels feature misconceptions about historical determinism, as well as a failure to acknowledge the relational aspects of causation (p.11; c.f. Coffin, 2006; Historical Association, 2019). Stages 3 to 5 begin to introduce ideas such as the inter-relationality of causes (i.e. net/web analogy), effective categorisation, and judgement and justification about relative importance (Scott, 1990, pp.11-12; c.f. Lomas, 1993; Coffin, 2006; Wineburg, 2007). Interestingly, whilst Scott (1990) identifies thematic categorisation and judgement as features of his Level 4, Lee and Shemilt (2009) criticise such an approach, stating that "[u]nless and until students operate at or beyond Level 6, factor classification and prioritisation tasks should be avoided… the categorisation of factors as 'political', 'economic', 'religious' and so on may amount to little more than the naming of relatively meaningless parts" (p.48). Scott's design as a progression model intended to guide teacher planning, compared to Lee's and Shemilt's evaluation of the qualities evident in samples of student work, makes the former more readily compatible with the pragmatic demand of seeking to improve students' conceptual understanding whilst meeting the requirements set by exam boards.

Figure 1: summaries of Lee and Shelmilt's (2009) and Scott's (1990) progression models



The figure originally presented here cannot be made freely available via ORA because of copyright. The figure was sourced at Lee, P. and Shemilt, D., (2009). Is any explanation better than none? Over-determined narratives, senseless agencies and one-way streets in students' learning about cause and consequence in history. *Teaching History* 137, pp.42-49; and Scott, J. (ed.) (1990). Understanding Cause and Effect: Learning and Teaching about causation and consequence in History. Teaching History Research Group, Harlow: Longman.

multiple ways of categorising, linking, and relating causes, with a general consensus emerging that the act of determining and justifying the "ultimate cause" (Carr, 1986, p.84) is the most challenging aspect, at least at school level (see Watts, 2016, p.157 for his views on how this differs at university). Even before arriving at the task of prioritisation, however, cause and consequence clearly operates as a highly complex concept, leading to numerous challenges for students (Haydn et al., 2015).

b. What do students find challenging when constructing causal arguments?

Causation is a multi-faceted concept that is central to the discipline of History. For any teacher to hope to support their students with understanding the concept, an understanding of the challenges that they are likely to face is essential.

Chronology and sequencing can prove challenging; students stand little chance of explaining why something happened, let alone what the most important reason was for its occurrence, if they lack secure knowledge of the order of events (Woodcock, 2010; Haydn et al, 2015; Chapman 2017; Historical Association, 2019). Rogers (2011) has argued that having a big picture of the past that is thought of "in terms of a complex, connected stream of time… is absolutely central to a fundamental

understanding of [History]" (p.51) and that this is something that many students lack, instead viewing the past as separate and unconnected.

Explanation in and of itself can prove challenging. Less competent students can display a tendency to simply list causes or factors and assume that doing so is sufficient to answer the question posed (Chapman and Woodcock, 2006; Worth, 2011; Morton and Seixas, 2012; Chapman, 2017). Failure to demonstrate an awareness or understanding of the effects of stated causes is another explanatory pitfall, resulting from treatment of causes as discreet things rather than as relationships *between* things (Shemilt, 1980; Woodcock, 2010; Haydn et al., 2015; Chapman, 2017). It has been argued that such difficulties arise from rote learning and a lack of deep knowledge (Chapman and Woodcock, 2006) as well as from how teachers talk about causation (Thompson, 1984). The complexity of the question can also contribute. Stanford (2019) gives an insightful explanation of this in his outline of his Key Stage 3 curriculum, beginning with the relatively simple Peasants' Revolt, where only consideration of the peasants' attitudes is required, moving to the dual roles of Charles I and Parliament in causing the English Civil War, and then to the multi-actor-based and multi-concept-based enquiry into the causes of the First World War. The more complex the question, the greater cognitive demands made of students. Finally, it must be remembered that processes such as establishing relative importance are cognitively demanding with "no easy template or algorithm"; many students find causal explanation hard because it *is* hard (Morton and Seixas, 2012, p.120).

Vocabulary and grammar can also present difficulties, as limitations in these areas can prevent students from fully communicated their ideas. This applies to language related to the substance of History, e.g. empire, serfdom, etc. and to the second-order concepts, such as cause and consequence (Counsell, 1997; Cooke, 2009; Haydn et al, 2015; Fordham, 2016). Coffin (2006) has extensively explored the challenges faced by students in this area, with sophisticated historical explanations requiring mastery of increasingly technical and abstract terminology. Indeed, the multi-faceted nature of causation requires the deployment of complex and wide-ranging terminology, so it is unsurprising that this is an area of challenge. Woodcock (2005) has shown how limitations in

students' vocabulary could result in "bland, unsophisticated, linear explanations (or worse, descriptions)" (p.6) due to students being ill-equipped to think about and then express their ideas. Carroll (2016a, 2016b and 2017) also posited that his students' poor understanding of complex grammar was holding back their ability to convey, or even develop, their ideas.

Much has also been written, particularly regarding Key Stage 3, about students' preconceptions and misconceptions. Students' tendency to treat past events as inevitable can result in their causal explanations being expressed scientifically or mechanistically, failing to take account of motive and historical probability (Lee and Shemilt, 2009; Morton and Seixas, 2012; Haydn et al., 2015; Chapman, 2017). Another misconception is viewing history as a monocausal linear chain, with one thing leading to another *ad infinitum* (Lee and Shemilt, 2009; Morton and Seixas, 2012). Similarly, many students tend to view causes as acting in isolation before converging together at the key event, a misconception that runs contrary to ideas about the "web-like" nature of causation (Woodcock, 2010; Haydn et al., 2015). Even if students do not hold such misconceptions, identifying and explaining these links can still pose a significant challenge (Morton and Seixas, 2012; McCrory, 2015; Haydn et al., 2015; Alcoe, 2015). Likewise, a lack of deep contextual knowledge can lead students towards making anachronistic judgements about people's behaviours and beliefs, in turn undermining their own analyses (Barton and Levstik, 2004; Woodcock, 2010; Haydn et al., 2015). Linked to this are the ideas expressed by Morton and Seixas (2012) in Guideposts 3 and 4 of their framework, where they explain that students with a poor understanding of causation fail to see the interplay between historical actors and their contexts, whilst also failing to grasp the unintended consequences of decisions. Students' thinking will often feature multiple misconceptions if they do not have a firm grasp of causation.

Whilst not an issue with causation itself, for students tasked with writing an extended piece of causal explanation, there are multiple cognitive demands at play which can affect the quality of their writing. Whilst the focus of this project is on causation itself, it is important to remember that issues such as memory and recall, the demands of how to structure an answer, and the interplay between

historical concepts can all present challenges when constructing an extended explanation (Counsell, 1997). Furthermore, the act of writing itself can be challenging for many students, especially when time pressure is introduced as a factor (Harris et al., 2011). Any intervention must take these myriad factors into account.

c. What does research suggest about how to help students improve their understanding of causation?

There exists a wide range of tools at the disposal of History teachers hoping to improve their students' understanding of causation. Whilst an attempt has been made to separate such methods into broad categories below, there is often some overlap; card sort activities intended to support categorisation may well also be used to help students think counterfactually, just as analogies can be used to develop students' historical vocabulary.

Recently, attention has been paid to the role of chronology and narrative and their links to improving students' understanding of causation (Woodcock, 2010; Rogers, 2011; Haydn et al, 2015; Chapman 2017; Historical Association, 2019). Chapman (2017) outlines one strategy beginning with a chronicle which can then be "storified" to add analysis and causal explanation, forming a web of causal relationships in doing so. Chapman (*Ibid*.) further developed this system with coding to show the different natures of the causes and their relationships. Rodker (2019) took a more direct approach to narrative, demonstrating with his Year 8 class that the production of non-linear narratives could be used to good effect to help students think about causal significance, link events, and gain confidence in their substantive knowledge. However, narrative must be used cautiously, to ensure that students' written work resembles a thematic causal analysis, rather than a story-focussed narrative.

Analogy has also been used to great effect to get students to think more conceptually about cause and consequence. Carr's (1986) story of the death of Robinson is a notable example, but perhaps the most famous, certainly amongst History teachers, is Chapman's (2003) tale of Alphonse the Camel, which he devised to enable students to think about typology and terminology (linking closely with

ideas about language and vocabulary, discussed below). The power of analogy is its capacity to make the unfamiliar familiar, thereby helping students to understand ideas that may have otherwise been beyond them (Fielding, 2015; McCarthy Young and Leinhardt, 2000). Examples include Chapman's (2003) lesson in which a student drew upon the game Buckaroo to help explain causation; similarly, Chapman's collaboration with Woodcock (2006) deployed a variant of the game Marble Maze. Analogies need not be so complex, however, as Buxton (2010) demonstrated when she encouraged students to weigh causes numerically, as have Morton and Seixas (2012) with their suggestion that students create their own analogies by explaining how they came to be in their present lesson. When using such interventions, however, it is imperative that teachers are aware of the limits of the models deployed and how far they accurately reflect the situation that is being analogised; debates around such issues have also been shown to help develop students' understanding (Woodcock, 2010; Chapman, 2017).

Interventions with a focus on language have also been shown to be useful for developing students' causal understanding, as they provide them with the tools for successful articulation (Morton and Sexias, 2012). As Woodcock (2005) is eager to point out, such interventions must stress the fact that language should not be used "for the sake of appearances" but rather as a means to "[enhance] students' historical analyses" (p.6); as he discovered through his own intervention, whilst enhancing students' linguistic capabilities can be valuable, it can also pose the risk that students become overly poetic and actually obscure their analysis. Carroll (2016b) went further, exploring how improving students' grammar can facilitate more effective conceptual understanding and communication (in this case, counterfactual argument). Carroll (2017) subsequently used group work with his Sixth Formers to help them use multi-voicedness whilst presenting counter arguments to a causal analysis of Hitler's rise to power.

Card sorts and other categorising activities often feature as a mainstay of causation-focussed lessons. A clear advantage of such exercises is their capacity to free students from the burden of potential mistakes; as nothing is committed to paper and the cards can easily be reconfigured,

students can experiment with confidence (Evans and Pate, 2007). Counsell (2004) has done significant work with card sorting activities, explaining that one benefit is how they allow students to visualise an emerging argument and the potential inadequacies of certain arrangements. Another strength of sorting activities is their multi-functionality; the same sets of cards can be used to get students to think about causes in different ways, whether by simply assigning them to different descriptive groups, arranging them chronologically, or organising them according to relative importance (e.g. Chapman (2003) with a diamond nine exercise; Evans and Pate (2007) with a classification activity; and McCrory's (2015) Venn diagrams in conjunction with other exercises to help think about links and relative importance). Evans and Pate (2007) have produced an excellent evaluation of card sorts, showing that, whilst there are many benefits to such an activity, there are also important pitfalls to avoid as well: namely a lack of appropriate teacher intervention, a failure to use card sorts in combination with an extended repertoire of tasks, whilst also ensuring that additional steps are purposeful, rather than deployed merely for the sake of it.

Diagrammatic visualisations have also been used by a range of practitioners to help students develop and elucidate their thinking, due to their power as "a fast route to identifying dynamics [and] understanding context" (Watts, 2016, p.155). Card sorts (discussed above) are one such tool in this category, but others also exist; Rogers' (2011) concerns about analogies such as barrels of gunpowder and camels and how they imply a formulaic, mechanical or strictly sequential nature to causation led him to develop a series of "causation maps" to demonstrate the overlapping variance of individual factors over time. He makes several helpful recommendations for their use, most notably in challenging students' misconceptions about historical inevitability, the scale of causation, and degrees of uncertainty in historical explanation. Alcoe (2015) and Bennett (2019) have both deployed link diagrams in their lessons, albeit in different ways and with different age groups (Sixth Form and Year 7, respectively), to get students to think more explicitly about the connections between causes and effects. Alcoe (2015) demonstrated the power of such tools to help students increase the speed at which they understand new ideas. As with many interventions, there is overlap here, such as with Chapman's (2017) chronology diagrams. McCrory (2015) also demonstrated the

power of diagrams when used in conjunction with other activities, supplementing her card sort activity with "why that size?" boxes to help students think about relative importance. The same goal lies behind activities such as Morton's and Seixas's (2012) relevance square and Counsell's (2004) zone of relevance, both of which can be used in conjunction with other card sorting tasks. One of the most powerful things about many of these activities is that, in visually depicting their thinking, students are creating an interpretation and standpoint, which will differ from those of their peers and which they will be forced to defend, developing their thinking in the process (Morton and Seixas, 2012; McCrory, 2015).

Counterfactual reasoning has received significant attention and attracted considerable debate, with critics cautioning against using it in fanciful ways (Megill, 2007; Evans 2014). It is therefore important to stress the need for any counterfactual exercise to be, as Buxton (2010) puts it "historically grounded" (p.7), a warning that has been made by historians (Ferguson, 2011) and researchers (Morton and Seixas, 2012) alike. Many of the aforementioned interventions have been used to encourage students to think counterfactually and challenge misconceptions about historical determinism and inevitability. Whilst Chapman (2003) simply peppered his lessons with "What ifs", Worth (2012) went to great lengths to design counterfactual resources, deploying a game about the fall of the Romanovs with a Year 10 class. Chapman and Woodcock (2006) collaborated to incorporate counterfactuals into a wider intervention featuring an adaptation of the game Marble Maze for their Year 11s, whilst Buxton (2010) used a comparative study of eighteenth-century Britain and France to help her Year 8s build a knowledge framework with which to make counterfactual assertions. Recently, Bennett (2019) demonstrated the value of using link diagrams and the simple yet effective question of "If X hadn't happened, would Y still have happened in the same way?" (p.21) to help Year 7 students comprehend alternate possibilities. Whilst acknowledging such a model to be imperfect, Bennett still found it an efficacious way of provoking counterfactual thought. Such practitioners, implementing a range of strategies across a wide range of year groups, have demonstrated the potential of counterfactuals to help students develop their causal reasoning, particularly when considering relative importance.

Interventions such as that developed by Worth (2012), which have value as counterfactual exercises, also make good use of the power of simulation or role-play. Asking students to adopt particular roles can be a valuable means of engrossing students in content and developing their understanding of human behaviour and agency, as well as the historical contexts within which such behaviour occurred (Luff, 2000; Moorhouse, 2008; Woodcock, 2010; Haydn et al, 2015). Linked closely with the concept of empathy, such interventions cause students to "give rise to hypotheses as to why [their] character chose, or was forced into, a particular course of action… [demonstrating that] motivation and reasoning are aspects of causation" (Luff, 2000, p.9). Such ideas were used in Worth's (2011) creation of a seventeenth-century village in which her Year 8 students were given characters, helping them understand the patterns of witchcraft accusation. Hamilton and McConnell (2003) advocate strongly for such approaches, stating that history without empathy reduces the subject "to a science" (p.19). Whilst not a role-play, Morton's and Seixas's (2012) diagrammatic activity, designed to get students to assess the interplay between individuals, groups, and contexts, supports the growth of similar understanding.

Models of extended writing are another commonplace way of supporting students. This seems logical, due to extended writing's centrality in historical explanation (Harris, 2001; Counsell, 2004). Carroll (2016a, 2016b and 2017) has deployed extracts by Richard Evans and Ian Kershaw in multiple ways to help enrich his Sixth Formers' writing, for example, using them to demonstrate how to express claim and counter-claim. Whilst most uses of historians' work as models or exemplars have been confined to A-Level students, Rodker (2019) has demonstrated their value with Key Stage 3 students, using extracts from Adam Tooze's *The Deluge* to help Year 8 construct more meaningful narratives and demonstrate causal significance. Whilst professional historians' work can be used as a gold-standard, teacher-produced model answers can also help students understand issues of structure and phrasing (Harris, 2001).

2. Self-regulation and metacognition

Students' ideas about themselves as individuals and learners matter because, by better understanding them, teachers can "develop and enhance the students' confidence in tackling challenging tasks, resilience in the face of error and failure, openness and willingness to share when interacting with peers, and pride in investing energy in actions that will lead to successful outcomes" (Hattie, 2012, p.40). As a result, there has been a vast amount of research and interest in the psychology of learning and how students' mindsets and metacognitive processes impact on their learning; far more has been written than could be adequately summarised here.[1] Furthermore, metacognition and self-regulation are multi-faceted and complex topics (Schunk and Ertmer, 2000; Wigfield et al., 2011; Goetz et al., 2013; Quigley et al., 2018) that are inextricably connected to other aspects of student mindset such as self-efficacy and motivation (Bandura, 1997; Bussey, 2011), as emphasised by Hattie's (2008) 'rope model' of self-concept, which illustrates the intertwined and interrelated nature of numerous different strands which impact students' performance and approach to tasks. Despite this, I have chosen to focus primarily on self-regulation in the following section, as this was the aspect that emerged as a focal point in my project. My school's current emphasis on it as an area of development for all teachers also played a role in determining this focus. It is also important to note that little direct research has been conducted on mindsets and metacognition specifically in History, let alone in a British context; therefore, much of what follows is drawn from international research on the topic in a more general sense.

[1] Some wide-ranging overviews, which have formed the basis of the following sections, and have proven to be useful starting points, include: Boekaerts et al. (eds.), 2000; Goetz and Hall (eds.), 2013; Hattie, 2012 and 2014; Quigley et al., 2018; and Schunk and Zimmerman (eds.), 2011.

a. What is self-regulation and why is it important?

Whilst definitions vary, Goetz et al. (2013) helpfully describe self-regulation as:

> *"a form of acquiring knowledge and skills in which the learners are independent and self-motivated. Learners independently choose their own goals and learning strategies that will lead to achieving those goals. [By] evaluating the effectiveness of one's learning strategies – comparing one's current state with the target state … learning can be modified and optimized"* (p.126).

Many researchers categorise self-regulatory behaviour using three distinct phases that centre around preparation for, completion of, and reflection on a given task (Schunk and Ertmer, 2000; Zimmerman, 2000; Pintrich and Zusho, 2002; Quigley et al., 2018; Muijs and Bokhove, 2020). Self-regulated learners will set their own goals and adopt strategies to best achieve them, monitoring their effectiveness and progress during the task and reflecting afterwards on their efficacy and success levels (Zimmerman, 2002). This **planning, monitoring, evaluating** triumvirate is a useful framework for thinking about barriers and strategies. The wide-ranging nature of self-regulation is shown by the other ways in which researchers have broken it down as a concept: Pintrich and Zusho (2002) divide it into cognition, motivation and affect, and behaviour. This schema is similar to that employed by the EEF in their recent report on self-regulation, where it is broken down into the three "essential components" of cognition, metacognition, and motivation, all of which support and interact with one another (Quigley et al., 2018, p.9). Consequently, self-regulation involves a wide range of processes, including but not limited to: setting goals for learning, deploying strategies for revision of material, monitoring performance, seeking assistance when required, holding productive beliefs about learning, reflecting and acting upon feedback, and maintaining motivation for optimal study (Schunk and Ertmer, 2000; Goetz et al., 2013). Ultimately, for any teacher wishing for their students to take greater responsibility for their own learning, self-regulation matters (Randi and Corno, 2000). The clear and widely-reported link between effective self-regulation and improved progress and student outcomes powerfully demonstrates its importance (Schunk, 1981 & 1984; Zimmerman and Martinez-Pons, 1986 & 1988; Zimmerman, 1990; Harris et al., 2011; Wigfield et al, 2011; Zimmerman and Schunk, 2011; Quigley et al., 2018).

As mentioned above, self-regulation links closely with other concepts such as student motivation and self-efficacy. High levels of student motivation contribute towards the selection of appropriate levels of challenge, persistence when confronted with difficulties or setbacks, and the application of deep learning strategies and effective self-regulation techniques (Dresel and Hall, 2013; Luftenegger et al. 2012). The *type* of motivation also matters: "the more the balance moves [from extrinsic] towards intrinsic motivation, the greater the investment in learning, which then leads to greater learning gains" (Hattie, 2012, p.42). More intrinsically motivated students are more likely to choose mastery goals, as opposed to performance- and socially-orientated goals (Grant and Dweck, 2003; Dweck, 2008; Hattie 2012). Hence, students exhibiting high levels of intrinsic motivation are more likely to engage with and deploy self-regulatory strategies and behaviours, improving their outcomes as a result. Self-efficacy is "the confidence or strength of belief that we have in ourselves that we can make our learning happen" (Hattie, 2012, p.41) and is **task-focussed**, varying from one situation to another (Bandura, 1986; Linnenbrink and Pintrich, 2003; Hattie, 2014). High levels of self-efficacy have been linked to higher levels of student performance and outcomes, significantly influencing effort levels, resilience in the face of difficulty or failure, and the ability to implement self-regulatory strategies (Zimmerman and Clearly, 2006; OECD, 2015). There is a close interplay between self-efficacy and self-regulation: individuals need to have confidence in their ability to self-regulate for it to be truly efficacious (Caprara et al., 2008). A range of studies, summarised in Schunk and Ertmer (2000), indicate a strong correlation between high levels of self-efficacy and self-regulation.

b. What are the barriers to effective self-regulation?

Self-regulation's complexity means that there are multiple barriers to its efficacy; if students lack any combination of the necessary understanding, tools and techniques, and motivation to effectively set goals, monitor their progress, and reflect on their performance and adapt their future approaches, then they will find self-regulation difficult (Goetz et al., 2013). As Quigley et al. (2018) state: "it is impossible to be metacognitive without having different cognitive strategies to hand and possessing the motivation and perseverance to tackle problems and apply these strategies" (p.9). Furthermore,

the very nature of schooling can prove a barrier; as Boekaerts and Niemivirta (2000) explain, “the generally accepted role pattern wherein teachers convey declarative and procedural knowledge and students must find a way to comprehend, store, and activate that knowledge leads to a situation in which students lack sufficient opportunity to organise and regulate their own learning” (p.417). The importance of autonomy is apparent: students can never hope to self-regulate if they are not provided with the freedom to determine aspects of their learning and are equipped with a range of strategies for doing so (Schunk and Ertmer, 2000). This bears striking resemblance to Hattie’s (2012) strand of self-dependence, where students become “dependent on adult directives” (p.44). If students are not clearly and explicitly shown how to self-regulate and are not encouraged to do so, with ample opportunity provided for practice, they are unlikely to develop the independent learning habits sought by teachers. Just as students need to have knowledge of strategies and techniques for regulating their learning, they also need to have secure subject knowledge, a complexity that comes from the relationship between the cognitive and metacognitive (Flavel, 1979; Pressley and Harris, 2006; Quigley et al, 2018). Students stand little chance of adequately self-regulating when writing an essay about the causes of Hitler’s rise to power if they lack both an understanding of what a good causal argument looks like and sufficient substantive knowledge about the topic.

Unsurprisingly, other aspects of student mindset contribute: a lack of self-efficacy belief about one’s ability to self-regulate and one’s abilities in general can hinder the process (Schunk and Ertmer, 2000). This can lead to a range of behaviours where students seek to avoid stressful situations that could damage their sense of self; this can take the form of misbehaving, daydreaming, procrastinating, setting avoidance goals (e.g. striving to avoid failure or to not do worse than one’s peers), self-discounting, dismissing feedback as not being valuable, developing learned hopelessness, and coming to believe that their ability or outcomes are not changeable (Pekrun, 2013; Hattie, 2014). Low self-efficacy can therefore contribute to inadequate self-regulation, impacting on learning. It has been suggested that levels of self-confidence are higher in boys than in girls of similar ability ranges (Hattie, 2014). Similarly, it has been suggested that women can be more likely to credit their success to external factors instead of their own competence, although this

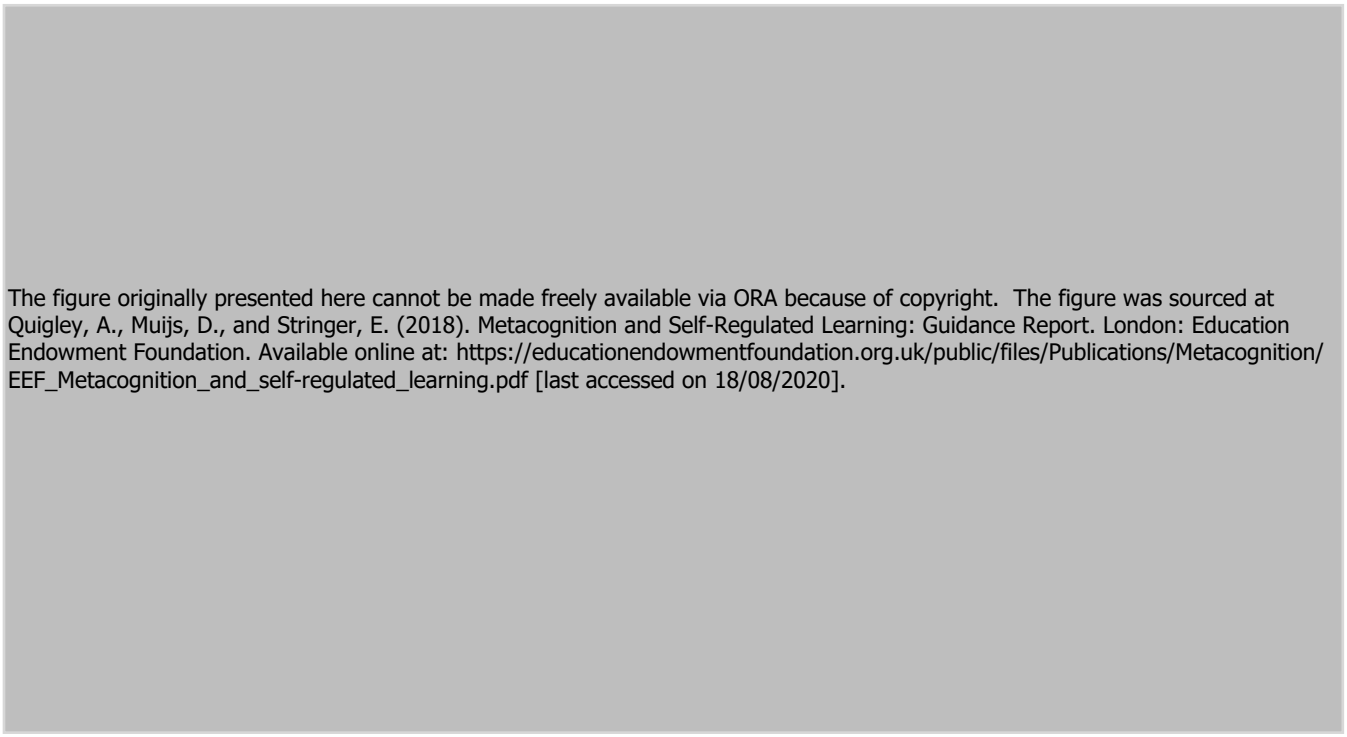
comes from research with a specific focus on mathematics and science, so extrapolations to History should be made cautiously (Dweck, 2008; Bussey, 2011). Girls' tendency towards perfectionism and a fear of mistakes may lead them to the erroneous conclusion that errors are a reflection on them as individuals; this, in turn, prevents them from engaging in trial-and-error or exploratory working (OECD, 2015). Similarly, Hattie's (2012) ideas around student self-dependence, which he states as applying particularly to gifted students, are pertinent; here, students tend towards extrinsic motivation and can become over-dependent on teacher directives to such a degree that they fail to develop the metacognitive skills needed for self-regulation. Such ideas are particularly relevant to my school's context.

c. What does research suggest about how to help improve student self-regulation?

Importantly, Hattie (2012) states that the metacognitive traits detailed above "can be learned" (p.46). Therefore, if one of the main barriers is students lacking understanding, tools and techniques, and motivation, then teachers seeking to improve their students' self-regulation should be able to do so by targeting these areas. As my primary focus is causation, my focus here is on techniques that can be embedded in regular teaching, rather than 'learning-to-learn' or study skills courses. Whilst these have been shown to be helpful (Weinstein et al., 2000), there is evidence to suggest that embedding interventions focussed on self-regulation within regular teaching can cause stronger effects on students as they are therefore better positioned to apply subject-specific techniques (Hattie et al., 1996; Randi and Corno, 2000; Seidel and Shavelson, 2007; Quigley et al., 2018). Despite this, as Muijs and Bokhove (2020) highlight, much of the existing research into the impact of self-regulatory interventions has been confined to Maths, Science, and English; hence, more general, rather than History-specific, recommendations make up much of what follows.

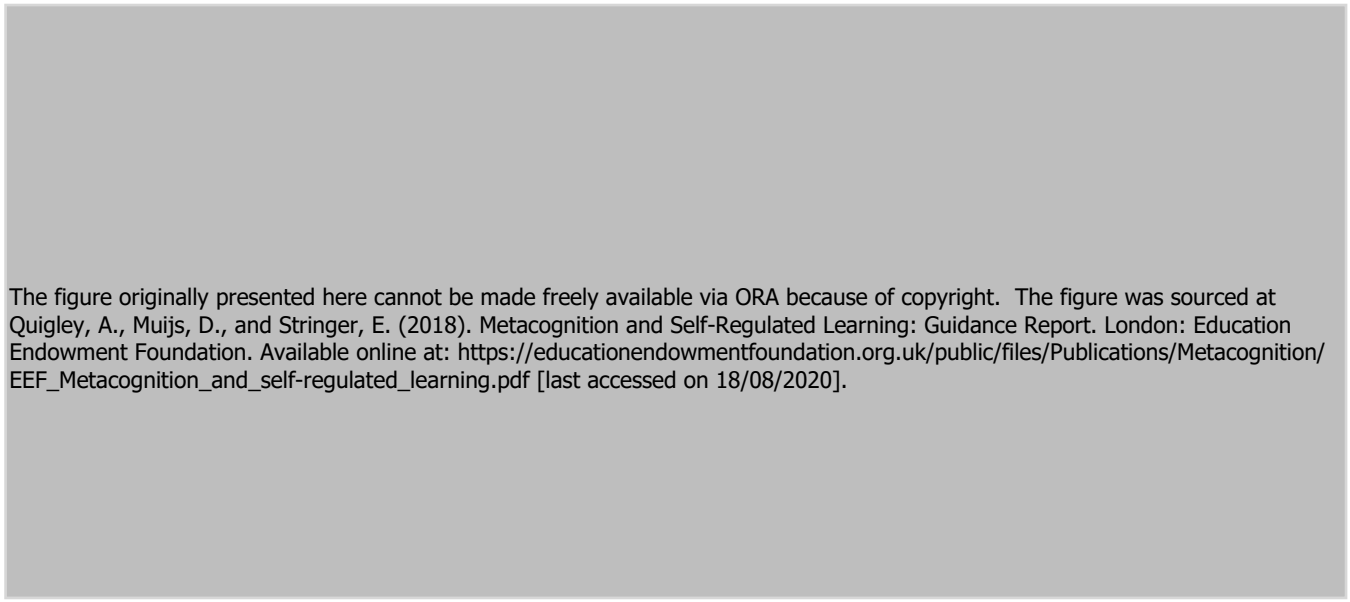
The EEF's *Guidance Report* (Quigley et al., 2018) on metacognition and self-regulated learning was produced with teachers in mind, drawing on a wide range of studies to provide seven recommendations (outlined in Figure 2, below) and is an accessible resource for approaching this issue.

Figure 2: EEF recommendations for developing metacognition and self-regulated learning (Quigley et al., 2018)



Whilst all seven recommendations are valuable, numbers 2 and 6 were most pertinent to my project in terms of designing classroom interventions with the aim of improving my students' causal understanding and self-regulation. The breakdown of approaches in the planning, monitoring, and evaluating phases (see Figure 3, below) presents a useful set of generic steps that teachers can take; the generic nature of the report, and the lack of History-specific research on self-regulation interventions are, however, limitations within the literature, meaning that it is up to individual teachers to devise the specifics of implementation (Quigley et al., 2018).

Figure 3: EEF metacognitive self-regulation strategies



The figure originally presented here cannot be made freely available via ORA because of copyright. The figure was sourced at Quigley, A., Muijs, D., and Stringer, E. (2018). Metacognition and Self-Regulated Learning: Guidance Report. London: Education Endowment Foundation. Available online at: https://educationendowmentfoundation.org.uk/public/files/Publications/Metacognition/EEF_Metacognition_and_self-regulated_learning.pdf [last accessed on 18/08/2020].

Quigley et al.'s (2018) choice of the word "explicit" is pertinent, as it is vital to help students see the value of developing their metacognitive processes (Hattie, 2012). Clarity of teacher explanation and demonstration are important for developing students' metacognitive and self-regulatory habits: this may take the form of clear and repeated discussion of goals (Harris et al., 2011; Hattie, 2012; Dresel and Hall, 2013), which should be mastery- and learning-focussed (Schunk and Ertmer, 2000) or direct explanation or modelling of self-regulatory techniques (Quigley et al, 2018). Teachers have an important role in making self-regulation and metacognition explicit, by integrating reference to them within their teaching, supporting and encouraging students to engage in appropriate processes, and fostering a classroom atmosphere that supports independent learning (Alton-Lee and Nuthall, 1990; Hattie, 2012). Teachers' language and demeanour are key aspects of this process; Hattie (2014, pp.224-226) provides a range of pointers about setting expectations, praising effort, stressing students' capacity to grow as leaners, and linking back to prior goals and achievements. Such measures can help to encourage students to 'buy-in' to productive mindsets and self-regulatory behaviour, an essential step for success (Schunk and Ertmer, 2000; Hattie, 2012).

Whilst the role of the teacher is important, it is also essential that students are given time for independent practice. Dresel and Hall (2013), summarising the work of Schiefele, have shown that, by providing students with a degree of autonomy, teachers can enable them to improve their self-

regulation and levels of intrinsic motivation. Such opportunities need not be complex and can include self-assessment, acting on meaningful feedback, and activities requiring a range of skills and degrees of independence (Dresel and Hall, 2013; Quigley et al, 2018). This independent practice needs to follow scaffolding and modelling by teachers, leading on to guided practice as students become more proficient (Randi and Corno, 2000; Harris et al, 2011; Goetz et al., 2013). With sufficient guidance, students should then be able to deploy strategies with greater independence (Harris et al., 2011; Quigley et al., 2018).

There is little existing literature on what interventions with a focus on self-regulation might look like in a History classroom. Quigley et al.'s (2018) recommendations, whilst useful, are generic and tend to focus on implementing explicit and clearly-focussed teacher talk and discussion about metacognition, as opposed to offering specific tasks or exercises that might be deployed or adapted. Similarly, Muijs's and Bokhove's (2020) distinction between cognitive strategies, aimed at mastering and selecting information and metacognitive strategies, aimed at supporting the planning, monitoring, and evaluation process, whilst useful, does not go into specifics about how each might manifest, particularly beyond the more-heavily researched subjects of English, Maths, and Science. Likewise, there is a limited amount of practitioner research in this area, although recent work by Richards (2019) on ways to provide students with planning tools is indicative of an increased awareness and interest in the topic amongst teachers.

Implications and Research Questions

The literature clearly indicated that causation is a complex concept for students to tackle, with prioritisation (i.e. creating an argument about relative importance) being one of the most challenging aspects of causal explanation. This difficulty is exacerbated by a range of barriers faced by students, including their own misconceptions and a lack of adequate vocabulary with which to express their understanding. A wide range of practitioners have experimented and found success with a variety of teaching interventions to tackle such issues, ranging from diagrams to analogy to vocabulary-targeting exercises. These findings prompted me to think more about my own students'

approach to causal arguments and revealed the need to gather data to develop my inclinations set out in my initial rationale.

Self-regulation has been shown to be an important concept that can make a marked difference to student outcomes. Covering planning, monitoring, and reflection, self-regulation encompasses a range of behaviours at different stages of completing any given task. Whilst students' own mindsets and the conditions in which they are taught can act as barriers to developing effective self-regulatory skills, these attributes can be learned; clearly explaining and deploying tools and techniques to do so in lessons and encouraging students to adopt them has been shown to be efficacious. As with causation, the ideas in the literature led me to learn more about my students' current mindsets and self-regulatory habits.

Therefore, the next stages of my research were defined by an attempt to answer the following research questions:

1. What existing understanding did my students have about causation and how was this exhibited in their responses to exam-style questions?
2. How far were my students able to self-regulate when tackling such questions?
3. Could a series of interventions that focussed on developing an understanding of causation and improving self-regulation improve my students' written work?

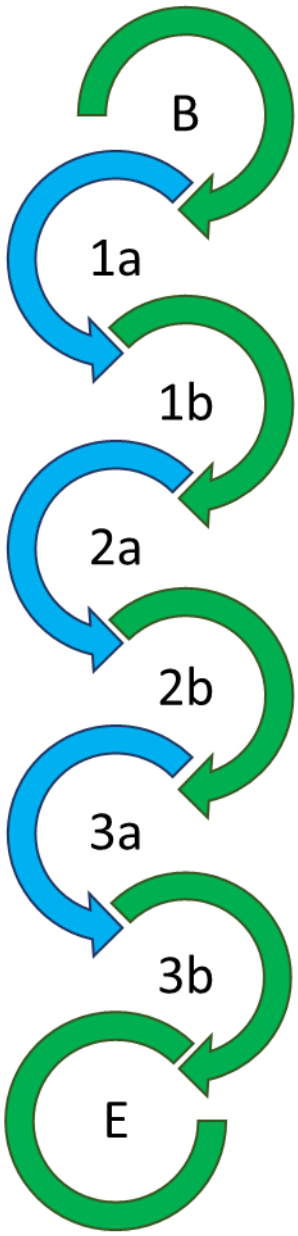
I sought to answer the first two questions in my baseline data collection, the analysis of which would then inform the first stage of my teaching interventions.

Methodology: overview

I used an action research model because, as Fielding (2015) suggests, interweaving conceptually-focussed tasks at intervals can provide students with time to develop their thinking. Furthermore, action research allowed me to revise and develop my interventions across the span of a year, responding to my students' emerging needs (Elliott, 1991). Finally, action research seemed appropriate, as it would allow me to approach causation's complexity incrementally, in line with Haydn et al.'s (2015) advice that causation "needs to be the focus of a series of lessons… [which] cover different aspects of understanding causation" (p.145).

The following sections are arranged sequentially, detailing the design of and findings from each phase, mirroring the research process. There were eight distinct stages; five focussed on collecting and evaluating data generated by students (green), whilst three were teaching interventions (blue).

Figure 4: Project overview



Baseline data: essays on reasons for the recovery of Weimar Germany, pupil survey, and think-out-loud activity.

Intervention 1; **link diagrams** to encourage pupils' conception of causation as "web-like".

Data collection and evaluation of Intervention 1: essays on why Hitler became Chancellor in January 1933.

Intervention 2; **significance circle diagrams** to develop arguments about relative importance.

Data collection and evaluation of intervention 2: essays on why the Cold War developed 1945-49 and how tension increased in the 1950s.

Intervention 3; **vocabulary-targeting exercises** to develop and enable expression of student understanding.

Data collection and evaluation of Intervention 3: pupils' mock exam responses and homework task.

Project evaluation, using pupil work from 3b and a second think-out-loud protocol and pupil survey.

NB. Following 2b, my school began remote learning due to COVID-19. The remaining phases were conducted digitally (implications discussed in relevant sections).

As I taught two Year 10 classes I chose to deploy my interventions with only one class. This was for comparison purposes, to help ascertain if my interventions produced a noticeable change in students' causal arguments and self-regulation (Hartas, 2010a). Whilst the non-intervention class did not represent a perfect control group due to differences between individual students, it would still provide a useful comparison point and give me greater confidence when determining if my interventions warranted repeated use. I discussed the ethics of this with my Head of Department

and supervisor; both agreed that there were no ethical obstacles, since the non-intervention class were not being disadvantaged, and the decision would aid my evaluation.

Baseline: initial data collection methodology and ethics

I began with an initial phase of data collection. I hoped to gain a deeper perspective than that outlined in my initial rationale, specifically in terms of my students' understanding of causation and their self-regulation.

Data was collected in three ways:

1. Through analysis of written work, using students' first causation-focussed essay.
2. Through "think-out-loud" exercises with a small sample of students, in which they were asked to plan their response to a different causation-focussed essay.
3. Through online surveys of all my Year 10 students. (See Appendix A).

As these three forms of data would form my baseline, I planned to utilise all three of them again at the end of my project to assess progress over the year, as well as to gauge students' perceptions of aspects of the process, giving my research a longitudinal design (Cohen et al., 2011; Hartas, 2010b; Cresswell, 2005). The decision to not conduct additional surveys and think-out-loud protocols after each intervention stemmed from a desire to limit time demands made of students and to let their written work form the primary source of evidence, as this was ultimately how their understanding would be assessed at the end of the GCSE.

The ethical considerations of my research are discussed throughout as appropriate. Details of the formal approval process can be found in my CUREC form (Appendix B).

a. Student essays

My students' first causation essay was: *"The main reason for the recovery of the Weimar Republic from 1923 to 1929 was the Dawes Plan." How far do you agree?* This follows the phrasing used in GCSE papers. I taught a sequence of lessons on the recovery of Germany in the 1920s and then planned the essay with my students, in my usual style. This was largely a structural exercise, in which

I managed discussion about the content of students' introductions, three PEAS (Point, Evidence, Analysis, Significance) paragraphs and conclusions. It was left to the students themselves to decide and justify their arguments. Central to this was the S section, where students justify the relative importance of the factor explained within the rest of the paragraph by linking and comparing it to the other factors.

I took a copy of my students' essays for analysis. Initially, I did not code the work, instead looking for similarities between my students' own work and the qualities and pitfalls highlighted in the literature. My analysis of students' work evolved over time, as discussed in subsequent methodology sections.

Importantly, there was nothing different about this sequence of lessons to my usual teaching and assessment at this point in the Year 10 course; the homework question was one of many such causation-focussed questions set throughout the year. This desire to ensure that my research did not distort methods of assessment or the workload placed upon my students, as well as not diminishing their exposure to and ability to practice questions on other concepts, meant that the overall structure of my research project was shaped around naturally-occurring causation topics within the Year 10 syllabus.

b. Think-out-loud exercise

Because my survey (discussed below) was mainly quantitative in nature, whilst my analysis of students' work did not allow for dialogue, I supplemented both with a more discursive form of data collection to provide a deeper insight into my students' thought processes and perspectives (Kvale, 1996; Hobson & Townsend, 2010). Therefore, several students were invited to participate in a think-out-loud protocol. Selection was based on academic data and students' engagement in lessons to cover a range of dispositions and capabilities. I chose to run the protocol with a sample due to the fact that such a method of data collection can be time-consuming, especially when organised within a school day with consideration for students' free time (Hobson and Townsend, 2010).

Participants were given a causation-based GCSE-style question and asked to plan it, verbalising their thinking process throughout. At times I would interject with questions to help clarify or develop statements (Leonard, 2003), but otherwise the flow of the exercise was determined by the students themselves, meaning that it represented a largely unstructured exercise (Hobson & Townsend, 2010) and represented a concurrent form of think-out-loud protocol (Kuusela & Paul, 2000). I hoped that this exercise would provide insights into my students' understanding of causation and well as their self-regulatory planning and monitoring. Students were asked to plan their answers with a stylus and iPad, with the device set to record both their writing and our conversation. This allowed me to fully focus on the exercise and evaluate the recording later, whilst also ensuring that data complied with my school's policies. Participants were made aware of the recording and its uses prior to agreeing to participate.

Following the advice offered by Hobson & Townsend (2010) and Cohen et al. (2011), I ensured that my students knew that participation was voluntary and that they could withdraw at any time. To make them feel at ease and attempt to offset any impact of the student-teacher power dynamic, I assured them that their responses would have no relationship to any formal assessment processes or how I worked with them in future, and that the exercise was solely for my own research purposes (BERA, 2018). Throughout the exercise, I made sure not to pass judgement on my students' comments, being mindful of my tone, body language and facial expression, whilst also ensuring that my questions were phrased to clarify participants' statements, rather than suggest answers.

c. Student surveys

The third component was a survey (Appendix A). I wanted to consider the views of *all* of my Year 10 students, alongside the deeper individual perspectives from the talk-through exercise, hence providing broad and deep insights into my students' understanding and attitudes (Cohen et al., 2011).

The survey was conducted digitally. This made ensuring anonymity easier, whilst also enabling rapid collection and presentation of findings, as this process was automated (Nesbary, 2000; Hartas,

2010b). As the host website logged the time of completion, I could distinguish one class from the other since they completed the survey in different lessons, allowing me to make comparisons at the end of the project. As my school regularly uses online surveys with students, the format was one with which my classes were familiar.

The survey was designed as a questionnaire, with mostly closed questions. These used a Likert scale, with students being asked how far they agreed with various statements (Hartas, 2010b). Three questions were open-ended, with students being asked to provide opinions on aspects of writing and arguing in History. The survey was primarily attitude-focussed to balance the more performance-focussed nature of the essays, which helped address the duality of my research and determine whether my interventions had an impact on both students' self-regulation and causal arguing (Cohen et al, 2011). Whilst I did not run a pilot survey, I asked other members of my department to read a draft, both to involve them in the process and to get their opinions on its utility.

Despite the survey's benefits, there were also potential pitfalls to consider, notably how some respondents may have been unwilling to be totally honest (Hartas, 2010b). To alleviate this, I made the survey anonymous and emphasised this to participants. Nevertheless, there remained a possibility that some students may have responded with what they *thought* I wanted to hear, rather than their honest opinions. Whilst anonymity somewhat mitigated this, it did prevent direct comparison of individual recipients' responses with my endpoint data. Although it would have been beneficial to be able to make this comparison, the logistical and technical limitations of designing a survey that would allow this whilst preserving anonymity extended beyond my resources and competencies; I mitigated this by comparing individuals via my other two sources of data.

The surveys were issued in lessons, via students' iPads. This allowed me to explain the purpose of the survey and the data (Hartas, 2010b). I emphasised that participation was voluntary and responses anonymous (BERA, 2018). The fact that one participant stopped answering part-way through suggests that the voluntary nature was understood. To ensure the data's integrity, I

explained that I could not help or offer clarification about questions or terminology. I also required the surveys to be completed individually and without discussion, to ensure that the results were reflective of each individual student.

Baseline data findings

a. Student essays

The responses were indicative of many of the pitfalls discussed in the literature. Many students made rudimentary and incongruous use of counterfactuals to try to argue their case for Stresemann being the main cause of the Weimar Republic's recovery; the response below being typical:

> "*the Dawes Plan was signed and started by Stressman [sic], so without him it would not have happened. Therefore, Stressman [sic] was the most helpful reason to help rebuild Germany [sic]."* [emphasis added]

Here, there is an attempt to demonstrate Stresemann's importance by envisioning an alternate timeline where he did not exist. The student's mistake, and the mistake of many others who took the same approach, is the highly deterministic nature of their argument (c.f. Chapman, 2017). For them, Stresemann and **only** Stresemann could have introduced the *rentenmark*, orchestrated the Dawes Plan, signed the Locarno Treaties, etc. This assertion does not survive scrutiny; in Stresemann's absence, another politician would likely have overseen a German recovery, albeit one that may have occurred differently. The argument is weakened by a lack of nuance.

Other iterations of the above argument demonstrated a linear causal process at work in students' minds:

> *"the work of Stresemann was the main reason that the Weimar Government recovered because … it was under him that the political treaties were signed... He also advocated the Dawes Plan. Because it was due to him that these other factors aided the Weimar's [sic] recovery … the government's recovery was mostly down to him."* [emphasis added]

The underlined sentence reveals a flaw in this student's understanding; for her, Stresemann is ultimately the main cause of the recovery because he began the process (c.f. Woodcock, 2010). This

is overly reductive and prevents any recognition or weighting of the contributions of the numerous economic and political factors to the Weimar Republic's recovery.

Another clear issue was students' use of assertions instead of explanations, as the following S section demonstrates:

> "*although these points [about the Dawes Plan] greatly helped Germany's economy, Stresemann played a bigger role in 'repairing' Germany and especially it's [sic] people, after the issues in the years 1918-1923."*

Whilst comparison is attempted, there is a total lack of linking between the factors. This was commonplace, with few students, despite it being explicitly mentioned in the planning lesson, ending their main paragraphs with S sections; even fewer directly linked or compared their factors in them. This effectively reduced such responses to three separate themed explanations, devoid of underlying argument. Elsewhere, the lack of nuanced vocabulary was also notable; not a single response used vocabulary to indicate varying impacts; all factors simply "caused recovery".

My analysis of students' work formed a useful first run of a process that I would repeat several times. It became clear that a coding system would be beneficial; this was something that I developed in subsequent phases to sharpen my analysis and facilitate comparisons between classes. I kept copies of all work submitted during the project, returning to them in my final evaluation.

<u>b. Think-out-loud</u>

This confirmed many of the findings from elsewhere, namely that, as Morton and Seixas (2012) say, students find causal argument difficult. Interestingly, several students changed their argument part way through the task. When queried, they all mentioned how getting into the detail and putting their ideas down on paper contributed to this, a revelation that chimed with ideas put forward by researchers such as Vygotsky (1986), Bakalis (2003), and Monte-Sano (2012). Although students appeared to have a good grasp of the factual content based on our discussion and note-taking from the lesson, plunging straight into answering a big causation question revealed a canyon in their

thinking; confronted with such a task almost immediately after merely covering the content suggested a need for supplementary tasks to bridge the gap.

Every student approached the task in terms of structure, beginning by outlining the need to have an introduction, three paragraphs, and a conclusion, with little thought for *what* they would fundamentally be about and what the argument would be. Having determined their factors, every student simply bullet-pointed a full answer, producing alternating lists of evidence and analysis. It was this process which seemed to support students in developing – and often changing – their arguments, an element of the plan which was often addressed last.

Whilst only a few key findings emerged, what they revealed was significant. Namely, my students viewed the planning process as a structural one which amounted to writing an essay in short form. This contrasted with what I felt a plan should be: a device to think about and identify the essay's core points and fundamental argument. This was not necessarily surprising, as my students were simply mirroring the planning process that I typically employed in lessons. This led me to reflect critically on my own practice and was key in my decision to reshape the planning process during my second intervention. The fact that several students changed their arguments mid-way through, whilst suggesting a degree of effective self-regulatory monitoring, also implied that they lacked the planning skills to avoid such a scenario. Whilst the changes were often beneficial to the argument, the fact that my students had to plan several paragraphs before making such a realisation was troubling and implied that they began the task without sufficient engagement with the question. The desire to find tools to encourage students to deeply engage with their historical knowledge and understanding underpinned several of my interventions.

c. Student surveys

The survey results produced interesting revelations about students' perceptions of their own capabilities and preferences, as well as their understanding of History as a discipline and the concept of causation. All 33 students completed the survey, although one from the non-intervention class only answered the first six questions. The results are summarised in Tables 1 and 2. For the three

open-ended questions, I categorised the responses. In these, most students gave one response to each question, but a few made multiple points, making the total number of responses higher than the number of respondents.

Table 1: Baseline survey responses (Q1-8 & Q10)

#	Question	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree	No. of students
1	Writing arguments in History is easy	1	25		7	-	33
2	There is a "right" argument to questions in History		-		33		
3	I prefer to be told by my teacher what to argue when answering questions in History	2	12		16	3	
4	I find it difficult to decide which factor is the most important when writing in History	2	17		13	1	
5	I often change my mind about my argument when writing in History	1	19		11	2	
6	I feel like I have a range of tools and strategies to help me build my arguments in History	7	23		3	-	
7	I am confident at answering questions that deal with causation (why something happened)	7	20		4	1	32
8	Causation questions require a different argument to non-causation questions	3	12	15	1	1	
10	I am confident at explaining what causation means in History	5	15		9	3	

Shaded cells represent where response options were not provided.

Table 2: Baseline survey responses (Q9, Q11 & Q12)

Q9: What kinds of justification could be used to argue for the most important reason something happened?		
"Not sure" / Don't know"		8
Simple statements about "evidence"		6
Statements implying the most important factor was a "trigger" for other events or the start of a "chain of events"		12
Explicit mention of comparison of factors		1
Simple statements about a factor "affecting lots of people"		8
Q11: What aspects of writing an essay with an argument (e.g. part c questions) do you find EASY in History?		
Statements about:	Selecting and deploying evidence	12
	Structuring and writing the essay	9
	Analysing evidence	2
	Constructing and justifying an argument	5
Answers irrelevant to the question		5
Q12: What aspects of writing an essay with an argument (e.g. part c questions) do you find DIFFICULT in History?		
Statements about:	Selecting and deploying evidence	7
	Structuring and writing the essay	7
	Analysing evidence	3
	Constructing and justifying an argument	10
Answers irrelevant to the question		3

Importantly, it has been shown that students tend to over-estimate their own competency (Falchikov and Boud, 1989) and that less capable students tend to have poorer judgement when self-assessing their abilities (Hacker et al, 2000). Such studies indicate that caution is required when analysing student survey data. Nonetheless, the survey results were still valuable for gaining insights into my students' own perceptions and trends within their feelings and preferences. Regardless of the data's overall accuracy, I hoped that my endpoint survey would indicate the impact of my interventions, ideally in increased levels of student confidence.

Regarding self-regulation, Q3 was highly revealing, where 14/33 respondents expressed a preference for being told what the "best" argument was. This seemed to support my perception that my students needed to be empowered to construct arguments on their own. Fortunately, Q2 revealed that, whilst many students preferred to be told the "best" argument, all of them were at least aware that there often existed multiple credible arguments.

Regarding causation and arguments, Q4 and Q5 were both revealing, with 58% indicating that they found ranking factors (i.e. constructing an argument) difficult. This was particularly notable when compared with Q1, where 79% agreed that writing an argument was easy. Q5 shed further light on the problem, implying that often students were not sure of their argument when they began writing a question. Not only did this suggest a lack of deep engagement with the question before commencing writing, but also hinted at the need for better self-regulation in the planning and writing stages. It also supported my findings from the think-out-loud protocols, indicating that the issues faced by the sample of students were shared by many of their peers.

Q9 (Table 2) was revealing regarding how my students approached prioritisation. Only one directly mentioned comparison of factors, an element that is central to constructing an argument in my school's GCSE specification and one that is ingrained in the PEAS structure (where the Significance section acts as a mini-conclusion where the factors being considered are compared and linked). The survey therefore revealed that most students had a limited understanding of comparison's importance. Also notable was the high number of responses which revealed common misconceptions about causation, particularly the idea of historical linearity (Lee and Shemilt, 2009; Woodcock, 2010; Morton and Seixas, 2012), with whatever began the "chain of events" being regarded as the most important. Multiple "don't know" responses also indicated that many students lacked satisfactory levels of understanding about how to craft an argument. This also contradicted the high levels of confidence about tools and strategies displayed in Q6 and revealed a shortcoming of my survey; Q6 would have provided me with much greater insight had it been supplemented with a follow-up question asking students to state *what* these tools and strategies were. This was one of several changes introduced to the endpoint survey. The responses to Q11 and Q12 (Table 2) further supported the conclusion that creating an argument was the most challenging aspect of writing for my students, as suggested by the mark schemes and literature (Scott, 1990; Chapman, 2003; Pearson Edexcel, 2016; etc.).

Q7-10 (Table 1) all addressed causation directly. Whilst 84% indicated high confidence in answering causation questions (Q7), this number fell to 63% in Q10, which measured confidence at explaining what causation *is*. The fact that 53% of students either had no opinion or disagreed with the statement that causation questions differed to other types further supported the idea that my students were not as confident as Q7 suggested. Furthermore, several students asked what causation was whilst completing the survey. The responses to Q9 also indicated limited understanding, with no answers going beyond mentioning "triggers" or "chains of events". It therefore seemed that my students' struggles with historical argument was based on a weak understanding of the concept of causation; how could they form convincing, persuasive arguments if they lacked a deep understanding of the inter-relationality of events?

Implications for teaching interventions

These insights led me to develop my first intervention. There were three clear areas relating to causation that merited focus:

1. Students' lack of concept-specific and explanation-aiding language
2. Students' limited understanding of causation and regular use of misconceptions
3. Limited comparison of factors and frequent assertions of relative importance.

Whilst I wanted to improve all three areas, and felt that they overlapped in many ways, I was wary of attempting too much at once, thereby preventing my students from making progress. Conversations with my supervisor confirmed that initially focussing on one area was advisable. Woodcock's (2005) warnings about introducing language before conceptual understanding is firmly embedded led me to leave a language-targeted intervention until later. I then chose to focus on understanding causation, largely as I hypothesised that addressing this first would form a natural progression towards targeting the issue of comparison, as suggested by the literature's statements about progression from description to explanation to evaluation/prioritisation (Scott, 1990; Pearson Edexcel, 2016). I hoped that, by developing students' causal understanding and providing them with tools to support the construction of their arguments, their metacognition and self-regulation would improve; as

Watts (2016) suggests, "the two need not be separated, and the one could reinforce the other" (p.154). I would more directly approach metacognition in my second intervention.

I chose to focus Intervention 1 on addressing the "chain of events" misconception, as this was the most prevalent issue in the baseline data and seemed to underpin some of the other issues. Through this, I hoped to:

- Enable students to make better links between causal factors more frequently, allowing them to argue about relative importance more convincingly.
- Encourage students to see causation as a complex issue that was open to debate, reducing their reliance on me to provide the "right answer".

1a: Task Design: Link Diagrams

The first appropriate topic for this intervention was Hitler's rise to power. Taking inspiration from Alcoe's (2015) and Bennett's (2019) use of link diagrams, I developed one to cover the years 1929-33 (see Appendix C). My design differed in two ways:

i. I arranged the boxes in chronological order to prevent students from making time-defying links, as Alcoe (2015) experienced. The literature (Woodcock, 2010; Haydn et al, 2015; etc.) highlighted chronology as an area of challenge and I hoped to alleviate this, allowing my students to focus on the task's core demands. Also, this change aimed to help students clearly visualise the build-up to the end event, depicting long-, medium-, and short-term factors at play. I deliberately included long-term points about the re-organisation of the NSDAP in the 1920s and the powers of the President, hoping that this would force students to abandon a "first item in the process" means of justification.

ii. Rather than students annotating the diagram directly with their causal explanations, I included a second sheet alongside the link diagram. This was to ensure that the link diagram did not become messy, reducing its utility for supporting essay-writing and revision.

I piloted the diagram with two colleagues to check that no key information was omitted and that a wide range of links were possible. This piloting process also contributed to the decision to ask students to colour-code the boxes *after* making the links; both colleagues felt that this would make the realisation of the interplay between different factors more powerful.

I hoped that completing the diagram would remedy students' misconceptions about the idea of a monocausal linear chain (Lee and Shemilt, 2009; Morton and Seixas, 2012) or for multiple causes to act in isolation before causing the key event (Woodcock, 2010; Haydn et al., 2015). Essentially, I hoped to cultivate an understanding of causation that was much more "web-like".

My intervention class spent 30 minutes of lesson time working on the link diagrams following a normal sequence of lessons on the Nazis' rise to power. They were then set an essay for homework: *"The main reason Hitler became Chancellor in 1933 was because of the Great Depression." How far do you agree?* My non-intervention class were taught the same lessons content-wise but were not given the link diagrams to complete. Instead, I planned the essay with them in my normal, structure-based style. The responses from both groups formed my second set of data.

1b: Intervention 1: analysis, findings, and reflections

Coding methodology

My baseline essay analysis revealed the need for a coding system, the first version of which was implemented here. Since the written qualities I was trying to observe were different to those stated in the mark scheme (Pearson Edexcel, 2016), I needed a separate system. Whilst much more rudimentary than that developed by Hammond (2014) to code different types of knowledge, it stemmed from a similar desire to measure a "deeper *quality*" of work (p.18), akin to what Woodcock (2005, p.6) describes as "bold, powerful, nuanced, precise analyses and explanation". Considering the qualities that I hoped the link diagrams would cultivate in students' work, I implemented the following code, examples of which can be seen in Appendices D and E:

- Sophisticated links between factors
- Sophisticated links within factors
- Links to support a convincing argument
- Appropriate causal language

Whilst the benchmark for "sophisticated" is subjective, I based this on my own knowledge of students and the standard of work typically produced at this stage in the course. I also used my colleagues' expertise to review my judgements. My findings were subsequently presented in a History Department meeting. This proved a valuable opportunity for me to gain their input into my anticipated next steps and was a process I repeated following each phase.

"Web-like" explanations

It was encouraging to see many students begin to make more sophisticated links within and between factors, for example:

> *"Many of von Papen's actions unintentionally helped the Nazis gain more support and the Government not being able to function played right into the Nazi's [sic] hands."*
>
> "*After Bruning, the new Chancellor, slashed benefits and raised taxes, desperate citizens turned away from the weak and squabbling government and looked to extremist parties such as the Nazis for a solution… and it was the appeal via the Nazi's [sic] propaganda and tactics which made them the better option."*

Many intervention class students' work contained sections of comparable quality, with written explanations of how various factors interacted with one another to cause Hitler to be appointed Chancellor. It seemed that the diagrams, for this question at least, had helped students to develop their understanding of causation and express its "web-like" nature. Whilst some of the highest attaining students in the non-intervention group conveyed similar ideas, such explanations were more prevalent from those who had completed the link diagram, supporting Alcoe's (2015) experience. This conclusion was reinforced by my Head of Department who, when shown a sample of intervention class essays, agreed that the level of understanding was better than what she would typically expect at that point in the year.

Causal arguments

Few students, however, used these links to craft an argument about relative importance. This was not entirely surprising, since the link diagrams did not explicitly target this. However, a small number did make promising starts at using the links between factors to prioritise them, for example:

> *"Although the Great Depression and Hitler/Nazi campaigning meant increased support, von Papen and von Schleicher's actions actually forced Hindenburg to appoint Hitler [as Chancellor]."*
>
> *"The Great Depression was not the main reason… because although it helped the Nazis gain support and put pressure on Hindenburg, it was not the direct cause. Hindenburg's actions directly caused it to happen in 1933."* [underlining student's own]

The first quote has the foundations for a very good piece of causal explanation, although the concept-specific language used could be more precise; von Papen's and von Schleicher's actions might be described as a "trigger" or "tipping point", whilst the second quote implies that the actions of Hindenburg (and von Papen and von Schleicher) *determined* the time of Hitler's appointment. Such ideas were implicit, indicating that the students lacked the vocabulary to fully express their ideas; this was unsurprising, given that this intervention did not target vocabulary. Woodcock's (2005) query "does the linguistic release the conceptual?" immediately came to mind, contributing to a greater focus on vocabulary in future phases, particularly Intervention 3. What was explicit, however, was the attempt at using the relationship between events to justify an argument.

Although few used causal relationships to justify their arguments, it was notable that ten of the 17 intervention class students argued that the actions of von Papen, von Schleicher, and Hindenburg were the main reason; although none used such terminology, all based their argument on their actions acting as the tipping point. Only six non-intervention class students, all high attaining, made similar arguments, with the rest prioritising the Great Depression due to it starting the process. This suggested that the link diagrams had helped address some of the intervention class students' misconceptions about causation. Worryingly, whilst this was true for mid- and high-attaining students, it was the lower-attaining ones whose work improved the least, suggesting that either they were not utilising the link diagrams during completion of their homework, implying limited self-

regulation in terms of tools and strategies adopted, or that the link diagrams were not a helpful resource for them.

Other findings and reflections

Reflecting on the lesson, I was surprised by how long students spent working through their link diagrams. Whilst I had planned to use time at the end of the lesson to colour code the boxes to support thinking about categorisation and the links between factors, there was no time for this. In future, having finished the Year 10 course well in advance of the summer holidays, I intend to spend more time working on the link diagrams in class (both here and elsewhere in the course), as per Alcoe (2015) and Bennett (2019), whose extended use of similar diagrams clearly increased their benefits.

Whilst my first intervention helped most of my students form better links between events, this had not resulted in better arguments. Causal language remained an area in need of further development, and self-regulation needed more direct attention. Bearing in mind Woodcock's (2005) methodology of introducing language at the end of his intervention, it seemed appropriate for the next phase to directly relative importance, whilst also explicitly equipping students with better self-regulatory tools for planning and monitoring their extended writing (Quigley et al, 2018).

2a: Task Design: Circle Diagrams

Again, my focus on causation determined the intervention's timing; the next appropriate topic was the development of the Cold War in the 1940s. Considering the gains that Intervention 1's link diagram had produced, I devised another for this topic (see Appendix F: also piloted with two colleagues), now deploying it as a stepping-stone. Inspired by McCrory (2015) and Richards (2019), I sought a way to make relative importance and self-regulation central to the planning process. The result was a combination of McCrory's (2015) 'why that size?' boxes and Richards' (2019) 'spiderplans', with the former addressing relative importance and the latter addressing students'

need for self-regulation-supporting planning and monitoring. The 'circle diagrams' as they came to be known, required students to do three things:

1. Draw three circles, each one representing a factor, with the sizes of each circle representing each factor's relative importance.

2. Explain *why* each circle was a different size to the other two. To encourage thinking about causation, students were asked to draw arrows between the circles, indicating causal relationships.

3. Beside each circle, bullet point three key pieces of evidence to include when writing the paragraph; as with Richards' (2019) 'spiderplans', students were instructed to tick each point whilst writing the essay, thereby monitoring their progress.

The above represented a significant number of new processes for my students, so I therefore chose to stagger their introduction across two separate "mini-interventions":

1. The significance circles were introduced following a sequence of lessons on why the Cold War developed in the 1940s. Beforehand, students had spent a 40-minute homework completing a link diagram. In the lesson, I talked students through an exemplar circle diagram that I produced, modelled my thinking process throughout (Quigley et al, 2018). Then, taking inspiration from Counsell's (2004) 'zone of relevance' and '1,2,3 change the question' activities[2], the rest of the lesson was spent on multiple iterations of the question *"The main reason for the Cold War developing by 1949 was…"* with the three factors changing each time. This allowed multiple attempts at practising the circle diagrams independently (Quigley et al, 2018). Students were encouraged to use their link diagrams to see what causal relationships they had identified. One of the questions was then set for homework.

[2] Both exercises are card sorts, where students select evidence in response to changing questions, either by moving cards through decreasing zones of relevance, or replacing them in favour of more appropriate information.

2. Having marked and returned the first essays, I then taught a sequence of lessons on the Cold War in the 1950s. I then introduced the key evidence bullet points to provide supporting detail to the plan, based on Richards' (2019) 'spider planning'. Students began by repeating the above process; they then added three bullet points beside each circle, listing the key points they wanted to make in each paragraph (see Appendix G for my exemplar). Students completed a plan for the question *"The main reason that tension increased between the Superpowers in the 1950s was because of the Korean War. How far do you agree?"* which was then set for homework. To encourage self-regulation, particularly in the monitoring phase, I instructed students to tick off each of their bullet points as they wrote, explicitly explaining that this was a self-regulatory strategy (Quigley et al, 2018).

I also wanted students to start thinking about different *types* of causal relationship and how to express them. Therefore, I introduced a distinction between 'necessary' (i.e. x was essential for y) and 'contributory' (i.e. z was not essential for y, but it helped it happen) causes, taking inspiration from some of the types described by the likes of Evans (1997) and Coffin (2006). I introduced only two words at this point out of a desire to keep the focus on relative importance, whilst also providing students with an initial way of categorising the relationships between events. This limited focus on language was built upon significantly in Intervention 3.

Both pieces of work were then coded in a similar style to Intervention 1 (see Appendix H).

2b: Intervention 2: analysis, findings, and reflections

The lesson which introduced the diagrams was a revelation; the circle diagrams energised students and stimulated discussion, making the planning process much more dynamic than with the non-intervention class. Students readily compared and challenged each other's plans, making excellent use of the new vocabulary to explain the causal relationships between events. Due to my school's focus on self-regulation, the Head of Professional Development observed the lesson, commenting that there was "superb" modelling and questioning which encouraged self-regulatory techniques.

Students' proficient use of the terms 'necessary' and 'contributory' to explain the relationships between various factors and the start of the Cold War distinguished work within the intervention class from that of the non-intervention class. However, an unintended consequence of introducing the necessary/contributory dichotomy was that many students developed a new misconception: that because one cause was necessary for other causes, it was therefore the main reason for the event in question. This is not always true, and such reasoning can reinforce the misconception that whatever event began the causal process is automatically the main reason. My reflections on the lesson and students' work revealed that an essential third component was missing: the idea of sufficiency. Much like how the Great Depression was necessary but not enough by itself for Hitler to be appointed Chancellor (an implicit idea in many essays from Intervention 1), arguably so too were the contrasting beliefs and personalities of Truman and Stalin. I challenged this misconception when returning students' work, and plan to modify this intervention in future by including a wider range of causal vocabulary, as present in Evans' (1997) and Coffin's (2006), amongst others', explorations of language.

Misconceptions aside, the quality of causal explanation and argument was noticeably better in the 1940s essays. This perhaps reflected the fact that students had completed a link diagram, which they did not for the second question. It could also reflect the questions themselves; the first question was more straightforward, requiring an explanation of how various events interacted to produce a specific outcome (the start of the Cold War). The 1950s question, however, required students to explain how various events, which did not always interact, contributed to an ongoing process (tensions increasing). Stanford's (2019) comments about the varying complexity of causal models across topics and questions immediately came to mind. Many intervention class students picked up on this difference themselves, complaining that the necessary/contributory dichotomy did not always allow them to link the factors and build a causal argument. This led me to reflect on how I might address these different types of questions in future years and further incentivised me to focus my final intervention on vocabulary.

Both sets of essays from the intervention class were accompanied by students' circle diagrams. With two exceptions, these were well-developed and detailed (see Appendix I). Only six of the 17 circle diagram plans for the second essay, however, had been used as monitoring tools during essay writing, with students ticking the bullet points. This was the last extended essay I was able to set before the end of the year, which unfortunately meant there were limited opportunities to further practice this technique. This highlighted the need to introduce it earlier in Year 10, providing more time to embed it in student practice (Quigley et al, 2018).

3a: Task Design: Vocabulary-targeting Exercises

Although no panacea, the circle diagrams had helped my students make progress in developing their arguments. Their utility as a planning tool meant that I continued to use them where relevant for the remainder of the year, although the only scope for this was in revision lessons, meaning that there was no opportunity for students to practise independently or tick off the bullet-points to regulate their essay writing. This, and my findings from Intervention 2, meant that my final intervention focused on language. The use of necessary/contributory in Intervention 2 had enabled students to approach causal arguments in a more eloquent and sophisticated way; however, these two words were not sufficient to express causation's complexity. Therefore, my third intervention addressed this by providing students with more new terminology. Introducing causal terminology at this late point was also influenced by Woodcock's (2005) work, where vocabulary was introduced at the final stage of a series of lessons, giving voice to implicit understanding and activating new ideas.

Models of extended writing, either by historians or teachers, can be powerful tools, demonstrating exemplar use of phrasing, terminology, and structure (Harris, 2001; Carroll, 2016a; Rodker, 2019). They therefore seemed suitable for introducing a wide range of vocabulary and consolidating the PEAS structure. Therefore, I developed the following tasks:

1. Students read and annotated an essay, which featured a wide range of sophisticated causal vocabulary, explaining the meaning and quality of various terms.

2. Using the first essay and their knowledge from the Year 10 course, students filled in a table, giving generic definitions and historical examples of a range of vocabulary.
3. For homework, students read a second essay, which featured no sophisticated causal vocabulary, editing it to include better causal terminology. (See Appendix J for all resources).

When selecting vocabulary, I drew on a range of examples from the literature (Lomas, 1993; Evans, 1997; Chapman, 2003 and 2017; Woodcock, 2005; Coffin, 2006) as well as my own knowledge of causation. The selection was by no means exhaustive, but one that I hoped would allow students to explain a range of relationships, whilst also addressing my misgivings about the necessary/contributory dichotomy: 'sufficient' was one of the new terms introduced.

Due to the proximity of end-of-year exams and my school's homework policies for this period, a homework essay could not be set. This, coupled with my desire to see if the skills and techniques fostered by all three interventions could be deployed under exam conditions, when self-regulation becomes especially important, meant that my students' exam responses were used to evaluate both Intervention 3 and the project as a whole. These responses were supplemented with a paragraph completed for homework in the week after exams, to help mitigate the impact of writing to time in a closed-book scenario and provide work more comparable to that from earlier phases.

3b: Intervention 3: analysis, findings, and reflections

The vocabulary exercise's impact appeared to be mixed. Of the 17 intervention class students, seven made no use of any sophisticated causal vocabulary in the exam, whilst eight (four the same as those from the exam) did not in the homework task. Most who made appropriately used vocabulary in the exam did so only occasionally, often using "trigger", "necessary" or "contributory" rather than any of the new terminology. Those who employed causal vocabulary in the homework task did so more frequently. This was surprising, since the homework task consisted of only one paragraph compared to the exam's extended essays, and likely reflects the absence of significant time pressure and the capacity for students to use their vocabulary tables during the task.

Table 3 – Student use of causal vocabulary

Intervention Class Student	Appropriate causal terms used		Non-Intervention Class Student	Appropriate causal terms used	
	Exam	Homework		Exam	Homework
A	1	-	A	1	-
B	1	4	B	1	-
C	1	-	C	-	1
D	-	1	D	-	-
E	-	-	E	-	-
F	1	2	F	-	-
G	-	-	G	-	-
H	1	1	H	-	-
I	-	1	I	-	-
J	1	2	J	-	-
K	-	-	K	-	-
L	2	4	L	-	-
M	2	-	M	-	-
N	6	-	N	-	-
O	-	-	O	-	-
P	2	6			
Q	-	3			

Whilst these results suggest that Intervention 3 had a limited impact, positives were evident. Comparison with the non-intervention class revealed that the intervention class made a much greater use of vocabulary than their peers; across both the exam and the homework task, only three non-intervention students used any sophisticated causal vocabulary, each on only one occasion, compared with the 13 intervention class students who made at least one use of appropriate vocabulary in at least one piece of work. The impact that this had on the quality of explanation was apparent when comparing work between the classes:

> *"the reason that the impact of the Munich Putsch is a more important factor than party reorganisation is because ultimately party reorganisation came as a result of the Munch Putsch…"* – non-intervention class student F
>
> *"the Munich Putsch created the underlying conditions that enabled the Nazi Party to survive… this is because it preconditioned party reorganisation and it also enabled Hitler to show his leadership and strong views."* – intervention class student P

Both quotes convey the same relationship between the Munich Putsch and the reorganisation of the Nazi Party; the latter, however, expresses it in a much more nuanced manner, indicating a deeper

level of causal understanding (although the construction of “preconditioned” as a verb could be substituted with more eloquent phrasing). This typified the contrast between students who made use of high-level vocabulary and those that did not.

When used, causal vocabulary was typically deployed with greater frequency in the homework task. Student B, for example, used four separate terms (compared to one in the exam), amongst them an explanation of how Hitler’s imprisonment “laid the foundations for increased Nazi support in later years” and “enabled [him] to reorganise the party.” Student L, who also made four uses, explained how party reorganisation gave the Nazis “latent strength” which was revealed by the Great Depression. Such students made effective use of vocabulary to convey a nuanced understanding of causation, raising the quality of their answers. Similarly, the one student to make more than two uses of causal vocabulary in the exam, doing so six times, elevated her explanation of NATO’s impact significantly, describing how it “laid the foundations for a more dangerous war [in the future]”. Interestingly, the same student, who deployed such varied language to excellent effect in the exam, made no use of it in the homework task, suggesting that further work and practice is needed to fully embed the vocabulary and to encourage self-regulatory monitoring to check for its use (Quigley et al., 2018).

Overall, Intervention 3 appeared to benefit some students, whilst seemingly having little impact on others. One explanation is that remote teaching limited discussion and engagement during the exercise, making it far more teacher-led than had it been conducted in school. Furthermore, the intervention occurred three weeks before the exams, so the vocabulary was unlikely to be fresh for all students, particularly those who did not revise it. Considering this, and feedback from students in the final think-out-loud protocols and surveys, I will adapt the intervention next year; this is discussed in my final section, below. Nevertheless, the evidence does support Woodcock’s (2005; p.11) conclusion that “language skills, as any historical skills, flourish with perseverance and directed support and guidance”.

E: Evaluation of project as a whole

Methodology

To evaluate the entire project, I again gathered data from three sources. The process mirrored that for collecting the baseline data with some small adjustments.

a. Student work

The end-of-year exam essays and homework paragraphs were also used here. The exam presented an excellent opportunity to see if my students could demonstrate the skills and understanding from earlier work under timed conditions, whilst also applying self-regulatory processes. However, the timed conditions added a new variable; instead of having a 40-minute homework to write an essay, students now had 25 minutes at most. Hence, the homework paragraph represented a piece of data that was more comparable to earlier work.

Notably, due to COVID-19, these exams were conducted at home. Whilst the need to adhere to timings and complete the exam as a "closed book" task was strongly emphasised, it is impossible to know if these instructions were followed, limiting the exams' integrity. However, since the qualities that I was looking for were metacognitive processes and complex conceptual understanding, it seems unlikely that students would have gained any significant advantages in these areas had they used their resources, especially as the exam essay had not been planned previously. Finally, the logistical and emotional impacts of remote learning will have affected different students in myriad ways, making my final data not wholly representative of my interventions' impacts under normal circumstances.

I compared the final pieces of work to students' baseline essays, as well as to the exam and homework pieces completed by the non-intervention class. The coding system from my previous two rounds formed a part of this analysis (see Appendix K), as did collaboration with my Head of Department to corroborate my judgements about the qualities in students' work.

b. Think-out-loud exercise

Unlike the baseline, these protocols were conducted remotely in one-to-one calls made with the software used for online lessons. As per my school's safeguarding policy, these conversations were recorded, providing me with a record for analysis. In the calls, I asked students to share their screen, and then to proceed with planning their answer to a causation question, verbalising their thinking process. I also asked several follow-up questions about the intervention tasks.

For continuity, the exercise involved the same six students as before, providing me with the chance to observe changes in individual participants that the anonymous surveys did not. When inviting the participants, I reiterated the exercise's voluntary nature, along with how their responses were not linked to any formal assessment processes or my future attitude to them (BERA, 2018). Of the six students, four responded to arrange a time and attended the call, limiting the number of comparisons I could make. Significantly, the two who did not participate represented the lower attaining and less motivated students in my classes, which impinged upon my ability to draw conclusions about such individuals and their progress.

c. Student surveys

The surveys were also conducted remotely. Since the original baseline surveys had been completed online in lessons, this difference is unlikely to have had a significant impact on engagement, as indicated by a 31/32 completion rate.

To create a comparative data set, the same questions were asked as before. However, additional questions were added (see Appendix L). Some asked students from both classes to evaluate changes in their competence and confidence since the start of the year. Additional new questions were asked of the intervention class, requiring them to evaluate the intervention activities. This combination of old and new questions allowed me both to make comparisons with the baseline data and to gain an insight into my students' feelings about their progress.

As with the baseline survey, students were assured of anonymity and that participation was voluntary. Whilst I made every effort to encourage them to be honest, it is impossible to rule out the possibility that some students may have provided answers reflecting what they *thought* I wanted them to say. Notably, the surveys were completed in the penultimate week of the academic year, soon after the return of exam results, both of which are likely to have influenced students' perceptions of their progress and competency.

Findings

a. Student work

The impact of the vocabulary exercises has already been discussed above. In my analysis of the overall project, I was looking for evidence of increased self-regulation through planning, as well as more sophisticated explanations and causal arguments.

Five of the 17 intervention class students completed a plan in their exam. These all made use of simplified circle diagrams. Although a minority, this compares favourably to the non-intervention class, where no-one completed a plan, tentatively suggesting that the circle diagrams have provided some students with an effective planning tool. It was also encouraging that two intervention-class students clearly monitored their use of vocabulary in the exam, crossing out basic phrases like "led to" and replacing them with more sophisticated choices such as "exacerbated". However, there was little more that the written work revealed about self-regulation; the think-out-loud protocols and surveys, discussed below, were more revealing.

Regarding causation and argument, some students produced impressive explanations in their exam responses, such as:

> "*NATO increased tension and enabled any future war to be more dangerous… NATO was significant as it laid the foundations for a more dangerous war, however, it was not as important as the Korean War because no actual fighting occurred."* – extract from intervention class student N's exam.

Interestingly, despite deploying sophisticated vocabulary and giving a nuanced causal analysis of NATO's impact, the above student did not use causation as a core part of her argument, which was instead more significance-based. The same was true of most students from both classes (although the intervention class typically did so to a higher standard). This likely reflects the question which, much like the second one from Intervention 2, asked about the main cause of a *process*, rather than of a singular event, in this case the development of the Cold War, 1949-56. Whilst some students made sophisticated causal links and deployed advanced vocabulary, few hit the heights of previous pieces of work. Both the question and the impact of writing under exam conditions, ostensibly without resources as support, probably contributed to this, suggesting that, whilst my various interventions have helped students improve their written work, more practice is required to ensure that this improvement is sustained in exam conditions.

The homework task gave a better account of the progress made by my intervention class, both compared to their baseline and their non-intervention class peers:

> "*[The political treaties were the most important] because with the people's trust the Weimar Republic was not in such a vulnerable position [and was] therefore able to take further steps towards recovery which, without the treaties, it wouldn't have been able to take."* – extract from intervention class student F's baseline essay.
>
> "*The Munich Putsch was a necessary cause for Mein Kampf [being written] which significantly helped the Nazis' survival… [The Munich Putsch] was a necessary cause of party reorganisation following its temporary ban…"* – extract from intervention class student F's endpoint paragraph.
>
> "*The Munich Putsch was the main reason the Nazi Party survived [as] Hitler used the trial to get national publicity and sympathy… His time in prison to reconsider how best to achieve power also led to party reorganisation."* – extract from non-intervention class student L's endpoint paragraph.

The second quote demonstrates several qualities from Scott's (1990) Stage 4, with the multiple ways in which the Munich Putsch impacted subsequent events explained and used to prioritise it as a reason for the Nazi Party's survival. Short- and long-term effects are explained, with appropriate language used. Whilst the idea of sufficiency could still be explored, the response represents a notable improvement from the start of the year and has several qualities not evident in responses

from the non-intervention class. Three students did not utilise sophisticated causal explanation or vocabulary in either their exam or homework task, suggesting that the interventions had not yet developed their competency to a point where it can be expressed in writing. However, the progress of the rest of the class indicated that, for most, the interventions had positive effects. Worryingly, the three students whose work showed little sign of development or self-regulation were those whose attainment placed them towards the lower end of the set; the need to make the interventions useful and accessible to all students is an area that I intend to reflect further on ahead of modifying and teaching them next year.

b. Think-out-loud exercises

Of the four participants, three used a circle diagram. The fourth presented her ideas in a more traditional, structure-based format, but with clear evidence of self-regulation in that she decided the ranking of her three factors before writing anything down. When asked about this, she explained the process as being akin to doing the circle diagram in her head, before using the plan as a scaffold for structuring the essay. As the survey data also suggests, this showed that the circle diagrams had been readily adopted by many in the intervention class. This was also supported by comments made by three different students during parents' evening, where they, unprompted, mentioned the circle diagrams as tools they found particularly useful.

In addition to using the circle diagrams, three participants explained that they planned more often than compared to the start of the year, and that the circle diagrams had facilitated this change, making the planning process quicker, easier, and more helpful when writing essays. The fourth, who did not use a circle diagram, explained that she liked using them to think about her argument, but that she preferred having a structure-based plan to shape her essay. This was encouraging evidence of self-regulation, as she was making informed decisions about which strategies to deploy for various situations, even if I hold doubts about the benefits of a structure-based plan.

When giving feedback on the link diagrams, circle diagrams, and vocabulary exercises, all students responded positively, confirming the exercises' value. Other comments helped me think about

future adaptations to the activities. Notably, two students mentioned how, although they found the vocabulary tasks useful, they had not had time to fully absorb all the terminology. Two students referred to how they lacked the time to do full circle diagrams in the exams, instead choosing to draw the circles without any of the supporting annotations. This indicated good self-regulation, but also made me aware of the need to be more explicit in my explanation of how to adapt the circle diagrams in different circumstances, with detailed ones for homework and revision, and fast, minimalist ones for timed conditions (Quigley et al, 2018).

c. Student surveys

The data below are coded by class (intervention class; non-intervention class). Numbers in square brackets represent the change *between* responses to the baseline and endpoint surveys. New questions are indicated with asterisks (*). Shaded cells indicate non-available response options.

From the intervention class, all 17 students completed both surveys; from the non-intervention class, 16 students completed the baseline survey (one stopped after Q6), and 14 students completed the endpoint survey (one student dropped History mid-year, another did not attempt it).

Table 4: Endpoint survey responses (Q1-6, 8-11 & 13-14)

#	Question	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
1	Writing arguments in History is easy	1 [-] 2 [2]	7 [-5] 10 [-3]		9 [5] 2 [-1]	- [-] - [-]
2	There is a "right" argument to questions in History		1 [1] 4 [4]		16 [-1] 10 [-6]	
3	I prefer to be told by my teacher what to argue when answering questions in History	4 [2] 7 [7]	8 [3] 5 [-2]		5 [-3] 2 [-6]	- [-2] - [-1]
4	I find it difficult to decide which factor is the most important when writing in History	2 [2] - [-2]	9 [-2] 5 [-1]		5 [-] 9 [1]	1 [-] - [-]
5	I often change my mind about my argument when writing in History	1 [-] - [-]	10 [1] 5 [-5]		5 [-] 8 [2]	1 [-1] 1 [1]
6	I feel like I have a range of tools and strategies to help me build my arguments in History *(See Table 5, below, for tools and strategies named)*	5 [-] 3 [1]	8 [-1] 8 [-6]		4 [1] 3 [3]	- [-] - [-]
8*	I have got better at constructing arguments in my History essays this year	4 7	11 7	2 -	- -	- -
9*	I have got better at constructing arguments independently in my History essays this year	2 3	12 8	2 3	1 -	- -
10	I am confident at answering questions that deal with causation (why something happened)	3 [-1] 1 [-2]	12 [2] 14 [4]		2 [-] - [-2]	- [-1] - [-]
11	Causation questions require a different argument to non-causation questions	1 [1] - [-3]	8 [-] 6 [2]	8 [-] 8 [1]	- [-] - [-1]	- [-1] - [-]
13	I am confident at explaining what causation means in History	2 [-2] - [-1]	13 [7] 8 [-1]		2 [-2] 6 [1]	- [-3] - [-1]
14*	My understanding of causation has improved this year	2 -	15 10		- 4	- -

Table 5: Endpoint survey responses (Q7)

Q7: Please list any tools and strategies you have to help you build arguments in History*		
Circle Diagrams	9	NA
Vocabulary resources	3	NA
S (Significance) Sections	2	2
Class resources (notes, textbooks, etc)	4	3
PEA (without mention of S sections)	1	3
Link Diagrams	1	NA
Plans completed in the lesson	1	2
Generic statements about how factors link together	-	1
Generic statements about the scale / longevity of factors	-	2
Responses that indicated a lack of understanding as to the meaning of “tools and strategies”	-	3

Table 6: Endpoint survey responses (Q12)

Q12: What kinds of justification could be used to argue for the most important reason something happened?		
Explicit mention of comparison of factors	-	1
Simple statements about a factor "affecting lots of people"	-	4
Use of counterfactuals	1	1
Statements about scale of impact	12	6
Statements about one factor being “necessary” or a “precondition” for others	6	-
Statements about causation, but without nuanced language (e.g. “it caused other events”)	6	5

Table 7: Endpoint survey responses (Q15-16)

Q15: What aspects of writing an essay with an argument (e.g. part c questions) do you find EASY in History?			
Statements about:	Selecting and deploying evidence	9	5
	Structuring and writing the essay	8	6
	Analysing evidence	4	3
	Constructing and justifying an argument	5	8
Answers irrelevant to the question		-	-
Q16: What aspects of writing an essay with an argument (e.g. part c questions) do you find DIFFICULT in History?			
Statements about:	Selecting and deploying evidence	5	7
	Structuring and writing the essay	2	2
	Analysing evidence	3	3
	Constructing and justifying an argument	12	8
	Timing	2	3
Answers irrelevant to the question		1	-

Table 8: Endpoint survey responses (Q17-21)

The following questions were only asked of the intervention class.

#	Question	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
17	I have found link diagrams useful for improving my understanding of causation	4	9	2	2	-
18	I have found circle diagrams useful for thinking about which factor is most important	5	10	1	1	-
19	I have found circle diagrams useful as a planning tool	5	9	2	1	-
20	I have found vocabulary tables and example essays useful for explaining my ideas	6	8	3	-	-

Q21: Please give any feedback (positive or negative) you have on the link diagrams, circle diagrams, and vocabulary exercises.		
Positive statements about:	Link diagrams	1
	Circle diagrams	8
	Vocabulary exercises	1
Negative statements about:	Link diagrams	-
	Circle diagrams	-
	Vocabulary exercises	1
Positive statements (no activity specified)		4
Negative statements (no activity specified)		1

Regarding self-regulation, the survey results were mixed. In Q3, a significantly higher number of students from both classes now agreed or strongly agreed that they preferred to be told what to argue. Encouragingly, in Q9, the vast majority also agreed or strongly agreed that they had improved over the year at independently constructing arguments, indicating that, whilst their preference may be to be told, they at least feel well-equipped to tackle such tasks themselves. Interestingly, the results were broadly similar across the two classes and in Q5, *more* non-intervention class students disagreed with the suggestion that they often changed their mind about their argument mid-essay. This could mean that I need to do more in future to explicitly address self-regulatory techniques and encourage my students to use them (Hattie, 2012; Quigley et al., 2018). Alternatively, it could reflect the fact that my interventions had made my students more aware of their thinking and changes to it.

The results from Q6 and Q7 showed more of a difference between the two classes; whilst the number of intervention-class students who felt they possessed tools and strategies for crafting arguments remained broadly the same, the number of non-intervention class students decreased. When asked in Q7 to name such tools and strategies, nine intervention class students identified the circle diagrams. Within the non-intervention class, the strategies mentioned were rudimentary (class notes, using PEA, etc.) and no strategy was identified by more than three students. These results supported the think-out-loud protocol about the circle diagrams being an effective tool in my students' eyes. This was confirmed by responses to the final four questions; whilst most intervention-class students agreed that all three exercises were useful, it was the circle diagrams that were singled out for positive praise in Q21, as exemplified by one student's comment:

> *"I found the circle diagrams particularly useful for developing a clear argument and deciding the most useful factor."*

Multiple comments praised the visual aspect of the circle diagrams and the speed at which they could be completed. Even though the various interventions appeared to have influenced students' written work at different rates and with different levels of consistency, it was reassuring that the vast majority saw them as worthwhile exercises, adding to the evidence of practitioners including Alcoe (2015), Bennett (2019), McCrory (2015), and Woodcock (2005).

Regarding causal arguments, it was surprising that more intervention class students than at the start of the year indicated that they were difficult. This potentially reflects that, now that they have a greater understanding of what it means to argue in History, my students appreciate that it *is* challenging (Scott, 1990; Morton and Seixas, 2012). Most intervention class students certainly still found constructing arguments challenging, with 12/17 of them mentioning it in Q16. Despite acknowledging its difficulty, most also agreed that they had improved at constructing arguments, both independently (Q9) and otherwise (Q8). This supported the conclusions from my other data points.

An improved understanding of causation among intervention class students was shown in their responses to Q13 and Q14; far more agreed or strongly agreed that they were confident at explaining what causation meant and that their understanding of the concept had improved. This comparatively better understanding amongst intervention class students was also evident in their responses to Q12, which asked them to list how a "most important" factor could be justified; six made use of specialist vocabulary, including "necessary" and "precondition", which no non-intervention class students did. Whilst not limited to causal arguments, it was encouraging to see 12 intervention class students give developed statements about scale. Whilst six non-intervention class students made comparable comments, most listed simple reasons that did not suggest a deep understanding of prioritisation. This difference was further supported by several non-intervention class students asking me to clarify what was meant by causation during completion of the survey.

A two-tailed t-test (Frost, 2020) indicated that there were several sets of responses where the differences were statistically significant (see Appendix M). Notably, applying the test to both classes' baseline data indicated no significant differences, demonstrating the two sets' initial comparability. Comparison of the endpoint responses revealed significant differences for Q1, Q13, and Q14, supporting my conclusions about the intervention class's improved understanding of the challenges of arguing in History and of causation as a concept. Whilst the test showed no significant differences when comparing the intervention class's baseline and endpoint responses, this is mitigated by the findings from the think-out-loud protocols and essays, which did demonstrate notable improvements. Furthermore, the test's indication of statistically significant differences in the non-intervention class's responses to Q2 and Q3, with more of them now thinking that there are "right" answers and preferring to be told what to argue, suggests that my interventions prevented such unhelpful attitudes from developing in my intervention class pupils and that my "normal" way of teaching contributes to pupils becoming overly dependent on me (Hattie, 2012).

Final reflections and next steps

Could a series of interventions focussed on developing an understanding of causation and improving self-regulation improve my students' written work? Although not always consistently shown, and with a high degree of variation between individual students, my interventions do appear to have had positive impacts, particularly in terms of providing students with a planning tool that places argument at the forefront of the process, adding weight to McCrory's (2015) findings. Similarly, the impact of the link diagrams and vocabulary resources on some students supports findings by Alcoe (2015) and Woodcock (2005), amongst others, that such interventions can be impactful. I therefore intend to continue using them and have shared the resources with my colleagues for use with their classes.

However, as with any activities taught for the first time, further refinements will be necessary, as acknowledged by other practitioners (e.g. McCrory, 2015). Some of my students' best work was produced following completion of the link diagrams, suggesting that they are valuable tools for tackling specific causal questions; I intend to develop more of these for appropriate points in the GCSE course, as well as for other key stages, mirroring Alcoe's (2015) and Bennett's (2019) implementation with sixth formers and Year 7 respectively. The impact of the circle diagrams is probably the most positive finding, and I plan to introduce them at the very start of Year 10 in future, as well as to the non-intervention class in Year 11. This will facilitate a greater impact on self-regulation, as students will have more practice with the bullet-point component and its use in the monitoring phase. There is also scope to adapt them for Key Stages 3 and 5, as it seems likely that the impacts that McCrory (2015) and I observed can be replicated elsewhere. Whilst the impact of the vocabulary exercise was not consistent, it did make a significant difference to the quality of explanation in some students' work, and it was encouraging to see a minority begin to refine and regulate their terminology. Next year, rather than introducing a swathe of vocabulary towards the end of Year 10, I plan to introduce appropriate terminology in smaller amounts throughout the year, as the introduction of "necessary" and "contributory" during Intervention 2 seemed to impact a

much larger proportion of students, becoming more firmly embedded in their thinking (despite the pitfalls associated with this dichotomy discussed above). Introducing concept-specific language at Key Stage 3 is also something that I am eager to pursue further with my colleagues, so that students arrive at GCSE with a better foundation in this area.

Self-regulation remains a focus for my school; whilst my research has had a positive impact on some students, empowering them to be more metacognitively engaged is a long-term process that cannot be accomplished during one year of History lessons. Building on my research, my school's Head of Professional Development has asked me to help lead CPD sessions on self-regulation and how to integrate it into regular classroom teaching; for this I will draw upon my literature review and my experience of integrating self-regulatory techniques within lessons, exploring with colleagues how this might be adapted for other subjects. I will also consider ways to improve other aspects of metacognition, such as self-efficacy and motivation.

This project has proven to be a worthwhile experience for me from a professional standpoint. Particularly, the merits of engaging with literature produced by other practitioners has been evident throughout, with my interventions taking significant inspiration from the likes of Alcoe (2015), McCrory (2015), Richards (2019), and Woodcock (2005). Learning from and adapting the tools and approaches of other History teachers has benefited my own understanding and teaching, as well as the progress of my students, and is something I will continue to draw upon in future, particularly in relation to my responsibility for developing the department's Key Stage 3 syllabus. Whilst the work of other teachers has proved invaluable, our experiences point to the need for larger scale work, capable of quantifying more systematically the impact of various interventions. More large-scale, quantitative analysis, similar to work conducted by Lee and Shemilt (2009) or a broader, quantified, and updated version of that done by Evans and Pate (2007), or an extension of the work done by Morton and Seixas (2012), with a specific focus on measuring the impact of various interventions, would be invaluable in this regard, as would further research into subject-specific techniques to improve student metacognition to supplement the more general advice provided by the EEF, whose

most recent publication (Muijs and Bokhove, 2020) cites only one history-specific study of metacognition.

Conducting action research has also been illuminating and is something I intend to replicate, albeit potentially on a smaller scale, in future. It has allowed me to develop my teaching throughout the year, and has also supported a re-evaluation of the marking process, leading me to use it to diagnose understanding and respond accordingly, rather than merely measuring attainment against exam board criteria (Hammond, 2014). Furthermore, the experience of conducting think-out-loud protocols has been illuminating as a method for gaining insights into student processes and attitudes, particularly when compared to the limitations of survey data in evaluating these areas. I intend to continue using such strategies to better understand my students and adapt my teaching to their needs.

Reference List

Alcoe, A., (2015). Post hoc ergo propter hoc? Using causation diagrams to empower sixth-form students in their historical thinking about cause and effect. *Teaching History 161*, pp. 16-24.

Alton-Lee, A.G. and Nuthall, G.A. (1990). Pupil experiences and pupil learning in the elementary classroom: An illustration of a generative methodology. *Teaching and Teacher Education: An International Journal of Research and Studies*, 6 (1), pp.27-46.

Ashby, R., Lee, P., and Dickinson, A. (1997). How children explain the 'why' of history: The CHATA research project on teaching history. *Social Education,* 61(1), pp.17-21.

Bakalis, M. (2003). Direct teaching of paragraph cohesion. *Teaching History 110*, pp.18-26.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Barton, K. and Levstik, L. (2004). *Teaching History for the Common Good.* Mahwah NJ: Lawrence Erlbaum Associates Publishers.

Bennett, T., (2019). What if there is another way? Year 7 use diagrammatic representations of counterfactuals to develop their causal reasoning. *Teaching History 174*, pp.18-24.

BERA (2018). *Ethical Guidelines for Educational Research, fourth edition.* Available online at: https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-2018 [last accessed on 19/08/2020].

Boekaerts M. and Niemivirta, M. (2000). Self-Regulated Learning: Finding a Balance Between Learning Goals and Ego-protective Goals, in M. Boekaerts, P.R. Pintrich and M. Zeidner, (eds.) *Handbook of Self-Regulation* (pp.417-450). London: Academic Press.

Burn., K., Carrier, J., and Fielding, A. (2019). The Role of Practitioner Research and Publication in Developing a Cumulative and Critical Tradition of Research in Classroom-Based Interventions in History Education, in G.J. Stylianides and A. Childs (eds.) *Classroom-based interventions across subject areas: research to understand what works in education* (pp.118-137). Abingdon: Routledge.

Bussey, K. (2011). The Influence of Gender on Students' Self-Regulated Learning and Performance, in D.H. Schunk and B. Zimmerman (eds.) *Handbook of Self-Regulation of Learning and Performance* (pp.426-441). Abingdon: Routledge.

Buxton, E. (2010). Fog over channel; continent accessible? Year 8 use counterfactual reasoning to explore place and social upheaval in eighteenth-century France and Britain. *Teaching History 140*, pp.4-15.

Caprara, G.V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G.M., Barbaranelli, C., and Bandura, A. (2008). Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. *Journal of Educational Psychology*, 100, pp.525-534.

Carr, E.H. (1986). *What is History?* 2nd edition, R.W. Davies (ed.) Basingstoke: Macmillan.

Carroll, J. (2016a). The whole point of the thing: how nominalisation might develop students' written causal arguments. *Teaching History 162*, pp.16-24.

Carroll, J. (2016b). Grammar. Nazis. Does the grammatical 'release the conceptual'? *Teaching History 163*, pp.8-16.

Carroll, J. (2017). 'I feel if I say this in my essay it's not going to be as strong': multi-voicedness, 'oral rehearsal' and Year 13 students' written arguments. *Teaching History 167*, pp.8-17.

Chapman, A. (2003). Camels, diamonds and counterfactuals: a model for teaching causal reasoning. in *Teaching History 112*, pp.46-53.

Chapman, A. (2017). Causal explanation, in I. Davies (ed.), *Debates in History Teaching*, *2nd Edition* (pp.130-143). Abingdon: Routledge.

Chapman, A. and Woodcock, J. (2006). Mussolini's missing marbles: simulating history at GCSE. *Teaching History 124*, pp.17-27.

Coffin, C. (2006). *Historical discourse: The language of time, cause and evaluation*. London; New York: Continuum.

Cohen, L., Manion, L., Morrison, K., and Bell, R. (2011). *Research methods in education*, *7th edition*. London: Routledge.

Cooke, S. (2009). *The Language Demands of History*. National Association for Language Development in the Curriculum, available online at: https://www.naldic.org.uk/Resources/NALDIC/Teaching%20and%20Learning/Documents/The%20language%20demands%20of%20History.pdf [last accessed on 18/08/2020].

Counsell, C. (1997). *Analytical and Discursive Writing at Key Stage 3*. Shaftesbury: Blackmoor Press.

Counsell, C. (2004). *History and literacy in Y7: building the lesson around the text*. London: John Murray.

Cresswell, J.W. (2005). *Educational Research*. Upper Saddles River, NJ: Pearson.

Dresel, M. and Hall, N.C. (2013). Motivation, in T. Goetz and N.C. Hall (eds.) *Emotion, Motivation, and Self-Regulation: A Handbook for Teachers* (pp.57-122). Bradford: Emerald Group Publishing Limited.

Dweck, C. (2008). *Mindset: The new psychology of success.* New York: Ballantine

Elliott, J. (1991). *Action research for educational change.* Milton Keynes: Open University Press.

Evans, R.J. (1997). *In Defence of History.* London: Granta Books.

Evans, R.J. (2014). *Altered Pasts: Counterfactuals in History*. London: Little Brown.

Evans, J. and Pate, G. (2007). Does scaffolding make them fall? Reflecting on strategies for developing causal argument in Years 8 and 11. *Teaching History 128*, pp.18-29.

Falchikov, N. and Boud, D. (1989). Student Self-Assessment in Higher Education: A Meta-Analysis. *Review of Educational Research*, 59(4), pp.395-430.

Ferguson, N. (2011). *Virtual History: Alternatives and Counterfactuals.* London: Penguin.

Fielding, A. (2015). Transforming Year 11's conceptual understanding of change. *Teaching History 158*, pp.28-37.

Flavell, J.H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive Developmental Inquiry. *American Psychologist, 34,* pp.906-911.

Fordham, M. (2016). Knowledge and Language: Being Historical with Substantive Concepts, in C. Counsell, K. Burn, and A. Chapman (eds.) *Masterclass in History Education: Transforming Teaching and Learning* (pp.43-57). London: Bloomsbury.

Frost, J. (2020). *One-Tailed and Two-Tailed Hypothesis Tests Explained*, blog post, available online at: https://statisticsbyjim.com/hypothesis-testing/one-tailed-two-tailed-hypothesis-tests/ [last accessed on 22/08/2020].

Goetz, T., Nett, N.E., and Hall, N.C. (2013). Self-Regulated Learning, in T. Goetz and N.C. Hall (eds.) *Emotion, Motivation, and Self-Regulation: A Handbook for Teachers* (pp.123-166). Bradford: Emerald Group Publishing Limited.

Grant, H., and Dweck, C. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology, 85*(3), pp.541-553.

Hacker, D. J., Bol, L., Horgan, D. D., and Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, 92(1), pp.160-170 .

Hamilton, A. and McConnell, T. (2003). Using this map and all your knowledge, become Bismark. *Teaching History 112*, pp.15-19.

Hammond, K. (2014). The knowledge that 'flavours' a claim: towards building and assessing historical knowledge on three scales. *Teaching History 157*, pp.18-24.

Harris, R. (2001). Why essay-writing remains central to learning history at AS Level. *Teaching History 103*, pp.13-16.

Harris, K.R., Graham., S., MacArthur, C.A., Reid, R., and Mason, L.H. (2011). Self-Regulated Learning Processes and Children's Writing, in D.H. Schunk and B. Zimmerman (eds.) *Handbook of Self-Regulation of Learning and Performance* (pp.187-202). Abingdon: Routledge.

Hartas, D. (2010a). Between-group Differences: Comparing Two or More Means, in D. Hartas (ed.) *Educational Research and Enquiry: Qualitative and Quantitative Approaches* (pp.352-377). London: Continuum.

Hartas, D., (2010b). Survey Research in Education, in D. Hartas (ed.) *Educational Research and Enquiry: Qualitative and Quantitative Approaches* (pp.257-269). London: Continuum.

Hattie, J., Biggs, J., and Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, 66(2), pp.99-136.

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

Hattie, J. (2012). *Visible Learning for Teachers.* London: Routledge.

Hattie, J. (2014). V*isible learning and the science of how we learn.* London: Routledge.

Haydn, T., Stephen, A., Arthur, J. and Hunt, M. (2015). *Learning to Teach History in the Secondary School: A companion to school experience*, 4th edition. Abingdon: Routledge.

Historical Association (2019). What's the Wisdom on… Causation? *Teaching History 175*, pp.24-27.

Hobson, A.J. and Townsend, A. (2010). Interviewing as Educational Research Method(s), in D. Hartas (ed.) *Educational Research and Enquiry: Qualitative and Quantitative Approaches* (pp.223-238). London: Continuum.

Kuusela, H., and Paul, P. (2000). A Comparison of Concurrent and Retrospective Verbal Protocol Analysis. *The American Journal of Psychology, 113*(3), pp.387-404.

Kvale, S. (1996). *Interviews: An Introduction to Qualitative Research Interviewing*. Thousand Oaks, CA: Sage.

Lee, P. and Shemilt, D., (2009). Is any explanation better than none? Over-determined narratives, senseless agencies and one-way streets in students' learning about cause and consequence in history. *Teaching History 137*, pp.42-49.

Lee, P. and Shemilt, D. (2011). The concept that dares not speak its name: Should empathy come out of the closet? *Teaching History 143*, pp.39-48.

Leonard, M. (2003). Interviews, in R.L. Miller and J.D. Brewer (eds.) *The A-Z of Social Research* (pp.167-171). Thousand Oaks, CA: Sage.

Linnenbrink, E.A. and Pintrich, P.R. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading & Writing Quarterly*, 19(2), pp.119-137.

Lomas, T. (1993). T*eaching and assessing historical understanding, revised edition.* London: Historical Association.

Luff, I. (2000). 'I've been in the Reichstag': rethinking roleplay. *Teaching History 100*, pp.8-17.

Lüftenegger, M., Schober, B., Van de Schoot, R., Wagner, P., Finsterwald, M., and Spiel, C. (2012). Lifelong learning as a goal: Do autonomy and self-regulation in school result in well prepared pupils?. *Learning and Instruction,* 22, pp.27-36.

McCarthy Young, K. and Leinhardt, G. (2000). Wildflowers, Sheep and Democracy: the role of analogy in the teaching and learning of history, in J.F. Voss and M. Carretero (eds.) *International Review of History Education, Vol 2* (pp.154-198). London: Woburn Press.

McCrory, C., (2015). The knowledge illusion: who is doing what thinking? *Teaching History 161*, pp.37-47.

McCullagh, C.B. (2004). *Putting Postmodernism in Perspective*. London: Routledge.

Megill (2007). *Historical Knowledge, Historical Error: a contemporary guide to historical practice.* Chicago: University of Chicago Press

Moorhouse, D. (2008). How to make historical simulations adaptable, engaging, and manageable. *Teaching History 133*, pp.10-16.

Monte-Sano, C., (2012). Toward Disciplinary Writing In History: Preparing The Next Generation. *Perspectives on History*, 50(5). Available online at https://www.historians.org/publications-and-directories/perspectives-on-history/may-2012/toward-disciplinary-writing-in-history [last accessed on 18/08/2020].

Morton, T. and Seixas, P. (2012). *The Big Six Historical Thinking Concepts.* Toronto: Nelson Education.

Muijs, D. and Bokhove, C. (2020). *Metacognition and Self-Regulation: Evidence Review*. London: Education Endowment Foundation. Available online at https://educationendowmentfoundation.org.uk/public/files/Metacognition_and_self-regulation_review.pdf [last accessed on 18/08/2020].

Nesbary, D. (2000). *Survey Research and the World Wide Web*. Upper Saddle River, NJ: Allyn & Bacon.

OECD (2015). *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. Paris: OECD Publishing.

Pearson Edexcel (2016). *International GCSE: History (9:1): Sample Assessment Materials*. Available online at: https://qualifications.pearson.com/content/dam/pdf/International%20GCSE/History/2017/specification-and-sample-assessments/international-gcse-history-sams.pdf [last accessed 19/08/2020].

Pearson Edexcel (2019). *International GCSE: History (9:1): June 2019 Paper 1.* Available online at: https://qualifications.pearson.com/content/dam/secure/silver/all-uk-and-international/international-gcse/history/2017/exam-materials/4HI1_01_que_20190604_@_pdf?626584066829331 [last accessed 19/08/2020].

Pekrun, R. (2013). Emotion, Motivation, and Self-Regulation: Common Principles and Future Directions, in T. Goetz and N.C. Hall (eds.) *Emotion, Motivation, and Self-Regulation: A Handbook for Teachers* (pp.167-188). Bradford: Emerald Group Publishing Limited.

Pintrich, P.R., and Zusho, A. (2002). The development of academic self-regulation: the role of cognitive and motivational factors, in A. Wigfield and J.S. Eccles (eds.) *Development of Achievement Motivation* (pp.249-284). San Diego: Academic Press.

Pressley, M. and Harris, K. R. (2006). Cognitive strategy instruction: from basic research to classroom instruction, in P. A. Alexander and P. Winne (eds.) *Handbook of Educational Psychology, Second Edition* (pp.265-286). Mahwah, NJ: Erlbaum.

Quigley, A., Muijs, D., and Stringer, E. (2018). *Metacognition and Self-Regulated Learning: Guidance Report.* London: Education Endowment Foundation. Available online at: https://educationendowmentfoundation.org.uk/public/files/Publications/Metacognition/EEF_Metacognition_and_self-regulated_learning.pdf [last accessed on 18/08/2020].

Randi, J., and Corno, L. (2000). Teacher Innovations in Self-Regulated Learning, in M. Boekaerts, P.R. Pintrich and M. Zeidner (eds.) *Handbook of Self-Regulation*. (pp.651-685). London: Academic Press.

Richards, H. (2019). Houston, we have a problem: how can we help all students structure written answers effectively and independently? U*npublished draft article.*

Rogers, R. (2011). 'Isn't the trigger the thing that sets the rest of it on fire?' Causation maps: emphasising chronology in causation exercises. *Teaching History 142*, pp.50-55.

Rodker, A. (2019). Austin's narrative: an exploratory case study, with Year 8, into what kinds of feedback help students produce better historical narratives of the interwar years. *Teaching History 174*, pp.8-15.

Schunk, D.H. (1981). Modeling and attributional feedback effects on children's achievement: A self-efficacy analysis. *Journal of Educational Psychology*, 74, pp.93-105.

Schunk, D.H. (1984). Sequential attributional feedback and children's achievement behaviors. *Journal of Educational Psychology*, 76, pp.1159-1169.

Schunk, D.H., and Ertmer, P.A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions, in M. Boekaerts, P.R. Pintrich and M. Zeidner, (eds.) *Handbook of Self-Regulation* (pp.631-649). London: Academic Press.

Scott, J. (ed.) (1990). *Understanding Cause and Effect: Learning and Teaching about causation and consequence in History*. Teaching History Research Group, Harlow: Longman.

Seidel, T., and Shavelson, R.J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Psychology*, 79(1), pp.57-68.

Shemilt, D. (1980). *History 13-16: Evaluation*. Edinburgh: Holmes McDougall.

Stanford, M. (2019). Did the Bretons break? Planning increasingly complex causal models at Key Stage 3. *Teaching History 175,* pp.8-14.

Thompson, D., (1984). Understanding the Past: Procedures and Content, in A.R. Dickinson, P.J. Lee and P.J. Rogers (eds.) *Learning History* (pp.168-186). London; Heinemann.

Vygotsky, L., (1986). *Thought and Language*. Cambridge, MA; MIT Press.

Watts, J. (2016). Causation, Chronology and Source Interpretation: Looking at School History from the Perspective of a University History Faculty: Reflections on Chapters by Ellen Buxton, Paula Worth and Rick Rogers, in C. Counsell, K. Burn. and A. Chapman (eds.) *Masterclass in History Education: Transforming Teaching and Learning* (pp.153-158). London: Bloomsbury.

Weinstein, C.E., Husman, J., and Dierking, D. (2000). Self-Regulation Interventions with a Focus on Learning Strategies, in M. Boekaerts, P.R. Pintrich and M. Zeidner (eds.) *Handbook of Self-Regulation* (pp.727-747). London: Academic Press.

Wigfield, A., Klauda, S.L., and Cambria, J. (2011). Influences on the Development of Academic Self-Regulatory Processes, in D.H. Schunk and B. Zimmerman (eds.) *Handbook of Self-Regulation of Learning and Performance* (pp.33-48). Abingdon: Routledge.

Wineburg, S. (2007). Unnatural and essential: the nature of historical thinking. *Teaching History 129*, pp.6-11.

Wolters, C.A., Benzon, M.B., ad Arroyo-Giner, C. (2011). Assessing Strategies for the Self-Regulation of Motivation, in D.H. Schunk and B. Zimmerman (eds.) *Handbook of Self-Regulation of Learning and Performance* (pp.298-312). Abingdon: Routledge.

Worth, P. (2011). Which women were executed for witchcraft? And which pupils cared? Low-attaining Year 8 use fiction to tackle three demons: extended reading, diversity and causation. *Teaching History 144*, pp.4-14.

Worth, P. (2012). Competition and counterfactuals without confusion: Year 10 play a game about the fall of the Tsarist empire to improve their causal reasoning. *Teaching History 149*, pp.28-35.

Woodcock, J., (2005). Does the linguistic release the conceptual? Helping Year 10 to improve their causal reasoning. *Teaching History 119*, pp.5-14.

Woodcock, J. (2010). Causal explanation, in I. Davies (ed.), *Debates in History Teaching* (pp.124-136). Abingdon: Routledge.

Zimmerman, B.J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25, pp.3-17.

Zimmerman, B.J. (2000). Attaining self-regulation: A social cognitive perspective, in M. Boekaerts, P.R. Pintrich and M. Zeidner, (eds.) *Handbook of Self-Regulation* (pp.13-39). London: Academic Press.

Zimmerman, B.J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory into Practice*, 41(2), pp.64-70.

Zimmerman, B.J. and Cleary, T.J. (2006). Adolescents' development of personal efficacy: The role of self-efficacy beliefs and self-regulatory skill, in F. Pajares and T. Urdan (eds.), *Self-efficacy Beliefs of Adolescents* (pp.45-69). Greenwich: Information Age.

Zimmerman, B.J., and Martinez-Pons, M. (1986). Development of a structured interview for assessing students' use of self-regulated learning strategies. *American Educational Research Journal*, 23, pp.614–628.

Zimmerman, B.J. and Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, 80, pp.284-290.

Zimmerman, B.J. and Schunk, D. (2011). Self-Regulated Learning and Performance: An Introduction and an Overview, in D.H. Schunk and B. Zimmerman (eds.) *Handbook of Self-Regulation of Learning and Performance*, (pp.1-12). Abingdon: Routledge.

APPENDIX A: Initial Student Survey

Page 1

* 1. Writing arguments in History is easy

- Strongly agree
- Agree
- Disagree
- Strongly disagree

* 2. There is a "right" argument to questions in History

- Agree
- Disagree

* 3. I prefer to be told by my teacher what to argue when answering questions in History

- Strongly agree
- Agree
- Disagree
- Strongly disagree

* 4. I find it difficult to decide which factor is the most important when writing in History

- Strongly agree
- Agree
- Disagree
- Strongly disagree

* 5. I often change my mind about my argument when writing in History

- Strongly agree
- Agree
- Disagree
- Strongly disagree

* 6. I feel like I have a range of tools and strategies to help me build my arguments in History

- Strongly agree
- Agree
- Disagree
- Strongly disagree

Page 2

Causation

* 7. I am confident at answering questions that deal with causation (why something happened)

- Strongly agree
- Agree
- Disagree
- Strongly disagree

* 8. Causation questions require a different argument to non-causation questions

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

* 9. What kinds of justification could be used to argue for the most important reason something happened?

* 10. I am confident at explaining what causation means in History

- Strongly agree
- Agree
- Disagree
- Strongly disagree

Difficulty

11. What aspects of writing an essay with an argument (e.g. part c questions) do you find EASY in History?

12. What aspects of writing an essay with an argument (e.g. part c questions) do you find DIFFICULT in History?

Appendix B: CUREC documentation

Confirmation of ethical approval

Dear [redacted]

TITLE: Developing Year 10 History students' written arguments through a focus on their self-regulation and decision-making
ED-CIA-20-012

The above application has been considered on behalf of the Departmental Research Ethics Committee (DREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly, approval has been granted.

Should there be any subsequent changes to the project which raise ethical issues not covered in the original application you should submit details to research.office@education.ox.ac.uk for consideration.

Good luck with your research study.

Best wishes
Nicola

Dr Nicola Warren-Lee
Geography PGCE Curriculum Tutor

University of Oxford/Department of Education
15 Norham Gardens, Oxford, OX2 6PY
Tel. 07970803012
Email nicola.warren-lee@education.ox.ac.uk
www.education.ox.ac.uk
@OxfordDeptofEd

Extracts from letter to Head seeking project approval

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION

15 Norham Gardens, Oxford OX2 6PY
Tel: +44(0)1865 274024 Fax: +44(0)1865 274027
general.enquiries@education.ox.ac.uk www.education.ox.ac.uk

Director Professor Jo-Anne Baird

14th October 2019

Dear

I am writing to enquire about conducting research in school this academic year. As you know, I am studying for the Master's in Learning and Teaching at Oxford University, supervised by Dr Katharine Burn. In my final research project *"Developing Year 10 History students' written arguments through a focus on their self-regulation and decision-making*", I will explore ways in which students can be encouraged to be more self-sufficient and self-confident when constructing written arguments.

I hope to conduct this research between October 2019 and July 2020, focussing on my two Year 10 classes. Much of the research will be designed to take place during the course of my normal teaching, making use of pupils' written work completed as part of their regular study. In addition, a selection of students (approximately 8 to 10), reflecting the ability spread of the GCSE cohort, will be asked to complete a (voluntary) exercise where they "think out loud" as they complete tasks requiring them to think about GCSE style essay questions, with the recordings of their reflections being taken for subsequent analysis by myself. Some questions may be asked of pupils during this exercise to help clarify and draw out their thinking. This aspect of the research will be conducted out of lesson time, during lunch time, so no students will be compelled to participate. This data collection will inform the development of a teaching intervention to be deployed with my Year 10 classes, with the aim of helping pupils plan their responses to written questions with greater confidence and independence, with a particular focus on formulating a line of argument. At the end of the process, participating students may be surveyed or interviewed in order to gather their thoughts. At various stages of the research, I will share my findings with other members of the History Department in order to draw on their insights and ideas.

Oxford University has strict ethical procedures on conducting ethical research, consistent with current British Educational Research Association guidelines. The University also recognises, however, that my study is a piece of practitioner research, and that schools already operate with the highest ethical standards. Therefore only <u>your</u> formal consent as headteacher is necessary, and not that of individual parents or staff. Although students will have to participate in the lessons delivered, as with any other lesson planned through the department, students and other teachers will be able to refuse to participate in any other research activities at any time.

All participants, including students, teacher and the school, would be made anonymous in all research reports. The original data would be kept strictly confidential, available only to my supervisor Dr Burn and me, and only used for academic purposes. Pseudonyms will be used

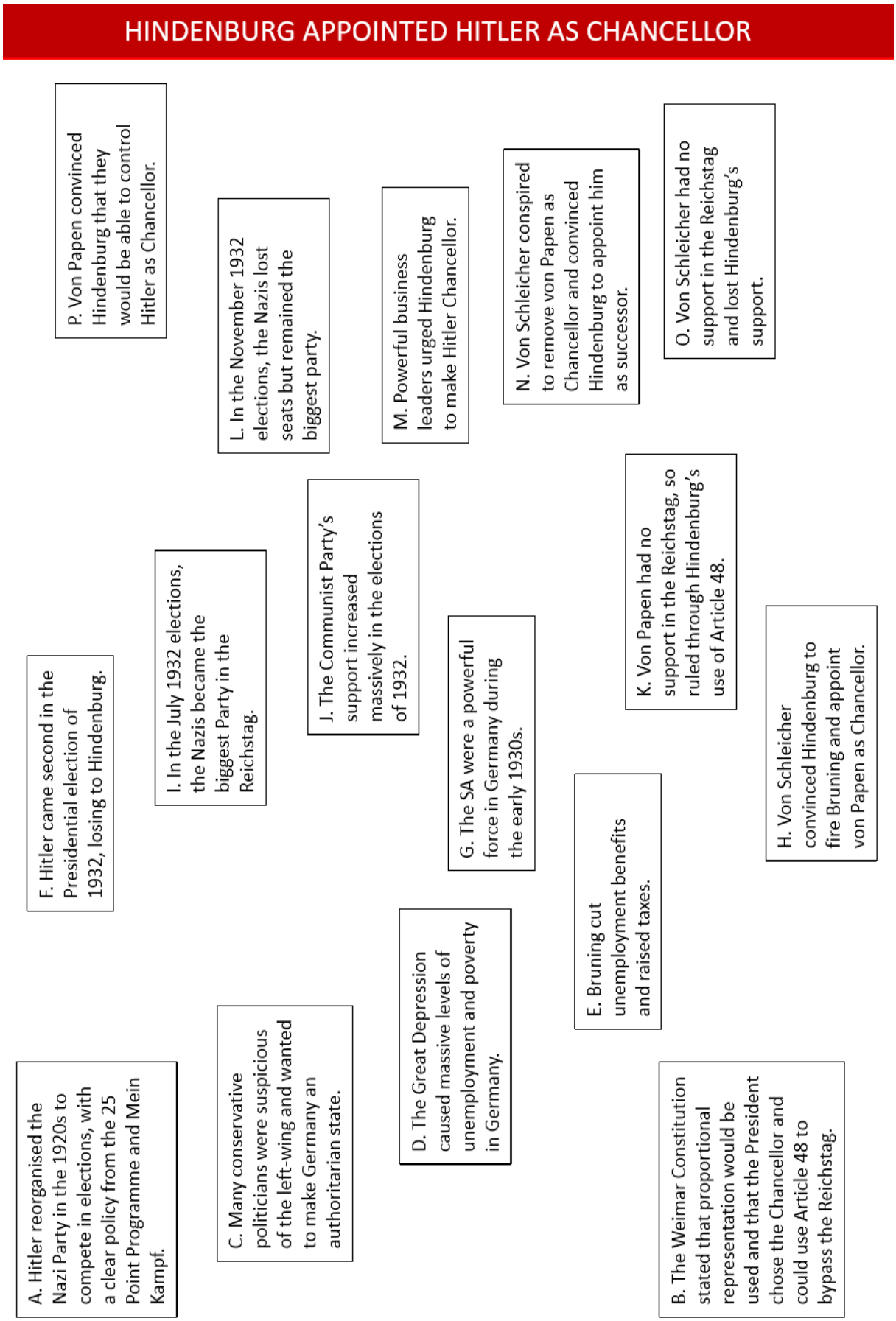
on all stored data rather than students' personal details. It will be kept for three years, in line with university guidelines before being deleted.

If you are happy for me to proceed with this study, please confirm that using the attached reply form. If you have any concerns or need more information about what is involved, please contact me or my supervisor (katharine.burn@education.ox.ac.uk). Further, if you have any questions about this ethics process at any time, please contact the chair of the department's research ethics committee, though: research.office@education.ox.ac.uk

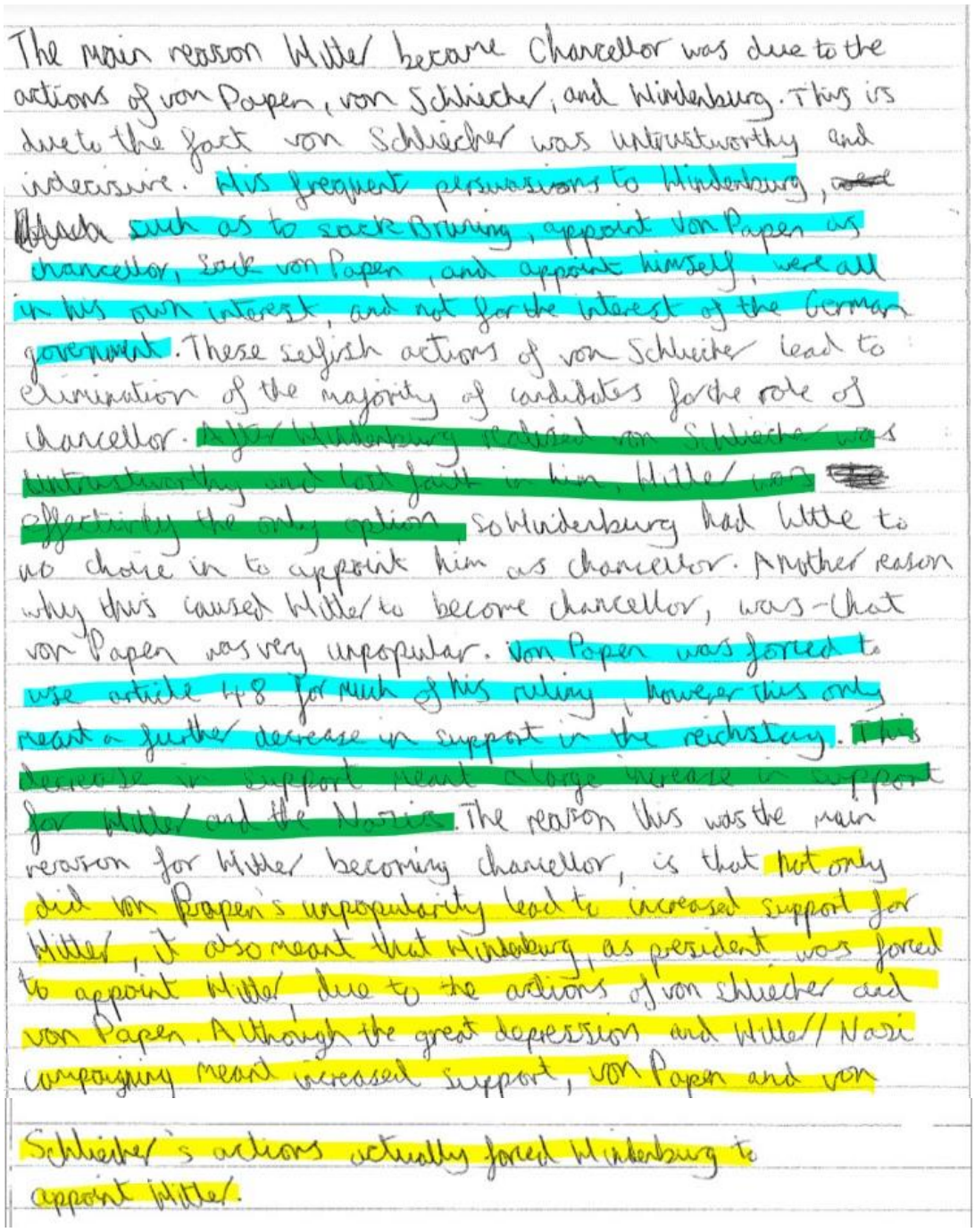
I look forward to hearing from you.

Yours sincerely,

APPENDIX C: Link Diagram template: Hitler's appointment as Chancellor



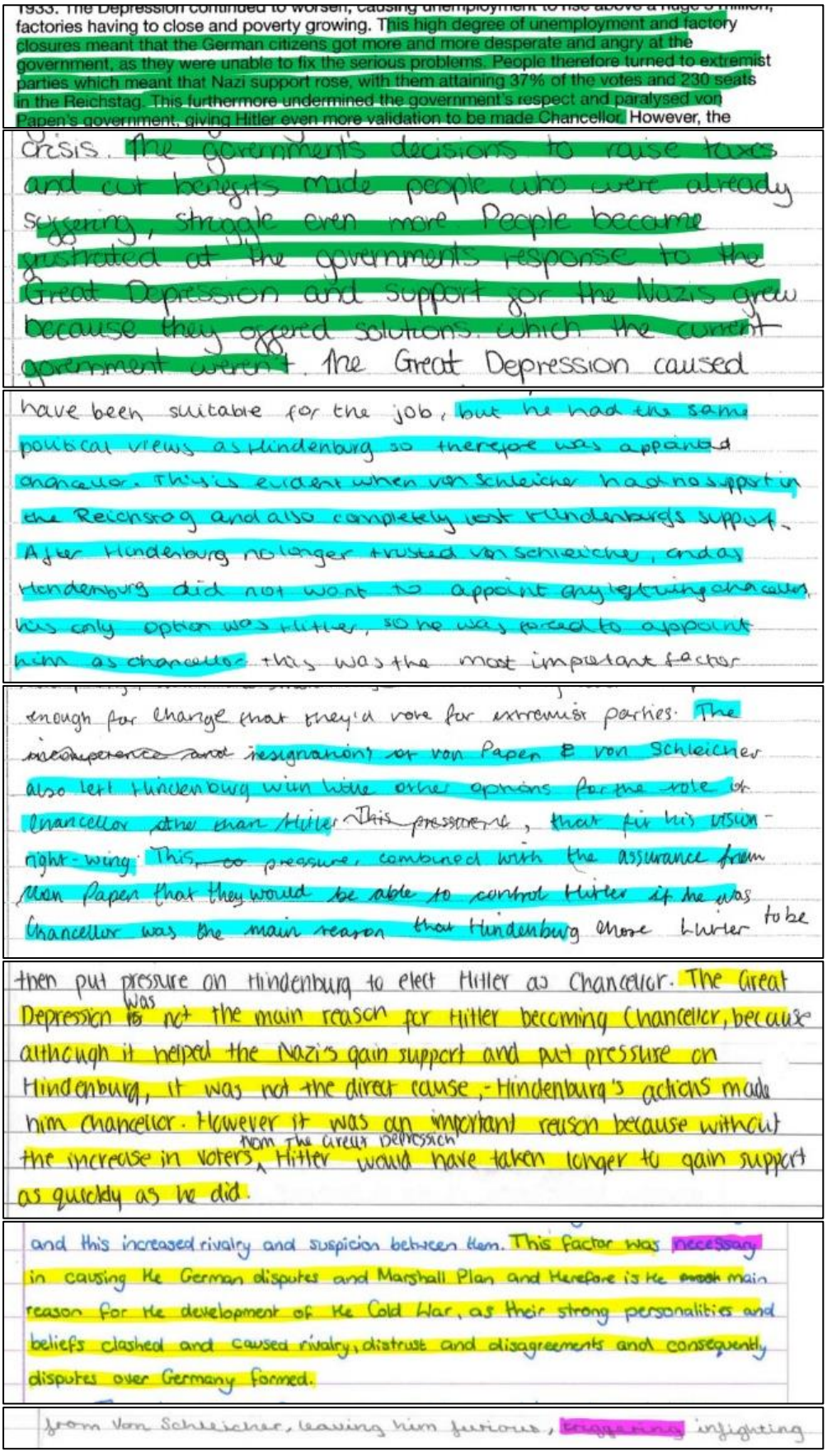
Appendix D: Example of coding of work from Intervention 1



The main reason Hitler became Chancellor was due to the actions of von Papen, von Schleicher, and Hindenburg. This is due to the fact von Schleicher was untrustworthy and indecisive. His frequent persuasions to Hindenburg, such as to sack Brüning, appoint von Papen as chancellor, sack von Papen, and appoint himself, were all in his own interest, and not for the interest of the German government. These selfish actions of von Schleicher lead to elimination of the majority of candidates for the role of chancellor. After Hindenburg realised von Schleicher was untrustworthy and lost faith in him, Hitler was effectively the only option so Hindenburg had little to no choice in to appoint him as chancellor. Another reason why this caused Hitler to become chancellor, was - that von Papen was very unpopular. Von Papen was forced to use article 48 for much of his ruling, however this only meant a further decrease in support in the reichstag. This decrease in support meant a large increase in support for Hitler and the Nazis. The reason this was the main reason for Hitler becoming chancellor, is that not only did von Papen's unpopularity lead to increased support for Hitler, it also meant that Hindenburg, as president was forced to appoint Hitler, due to the actions of von Schleicher and von Papen. Although the great depression and Hitler/Nazi campaigning meant increased support, von Papen and von Schleicher's actions actually forced Hindenburg to appoint Hitler.

Sophisticated links between factors	Sophisticated links within factors	Links used to support a convincing argument	Appropriate causal language

Appendix E: Illustrative examples of coding of pupil work



1933. The Depression continued to worsen, causing unemployment to rise above a huge 5 million, factories having to close and poverty growing. This high degree of unemployment and factory closures meant that the German citizens got more and more desperate and angry at the government, as they were unable to fix the serious problems. People therefore turned to extremist parties which meant that Nazi support rose, with them attaining 37% of the votes and 230 seats in the Reichstag. This furthermore undermined the government's respect and paralysed von Papen's government, giving Hitler even more validation to be made Chancellor. However, the

crisis. The government's decisions to raise taxes and cut benefits made people who were already suffering, struggle even more. People became frustrated at the governments response to the Great Depression and support for the Nazis grew because they offered solutions which the current government weren't. The Great Depression caused

have been suitable for the job, but he had the same political views as Hindenburg so therefore was appointed chancellor. This is evident when von Schleicher had no support in the Reichstag and also completely lost Hindenburgs support. After Hindenburg no longer trusted von Schleicher, and as Hindenburg did not want to appoint any left wing chancellor, his only option was Hitler, so he was forced to appoint him as chancellor- this was the most important factor

enough for change that they'd vote for extremist parties. The ~~incompetence and~~ resignations of von Papen & von Schleicher also left Hindenburg with little other options for the role of chancellor other than Hitler. This pressured, that fit his vision - right-wing. This pressure, combined with the assurance from Von Papen that they would be able to control Hitler if he was Chancellor was the main reason that Hindenburg chose Hitler to be

then put pressure on Hindenburg to elect Hitler as Chancellor. The Great Depression is was not the main reason for Hitler becoming Chancellor, because although it helped the Nazi's gain support and put pressure on Hindenburg, it was not the direct cause, - Hindenburg's actions made him chancellor. However it was an important reason because without the increase in voters, from the Great Depression Hitler would have taken longer to gain support as quickly as he did.

and this increased rivalry and suspicion between them. This factor was necessary in causing the German disputes and Marshall Plan and therefore is the ~~most~~ main reason for the development of the Cold War, as their strong personalities and beliefs clashed and caused rivalry, distrust and disagreements and consequently disputes over Germany formed.

from Von Schleicher, leaving him furious, ~~triggering~~ infighting

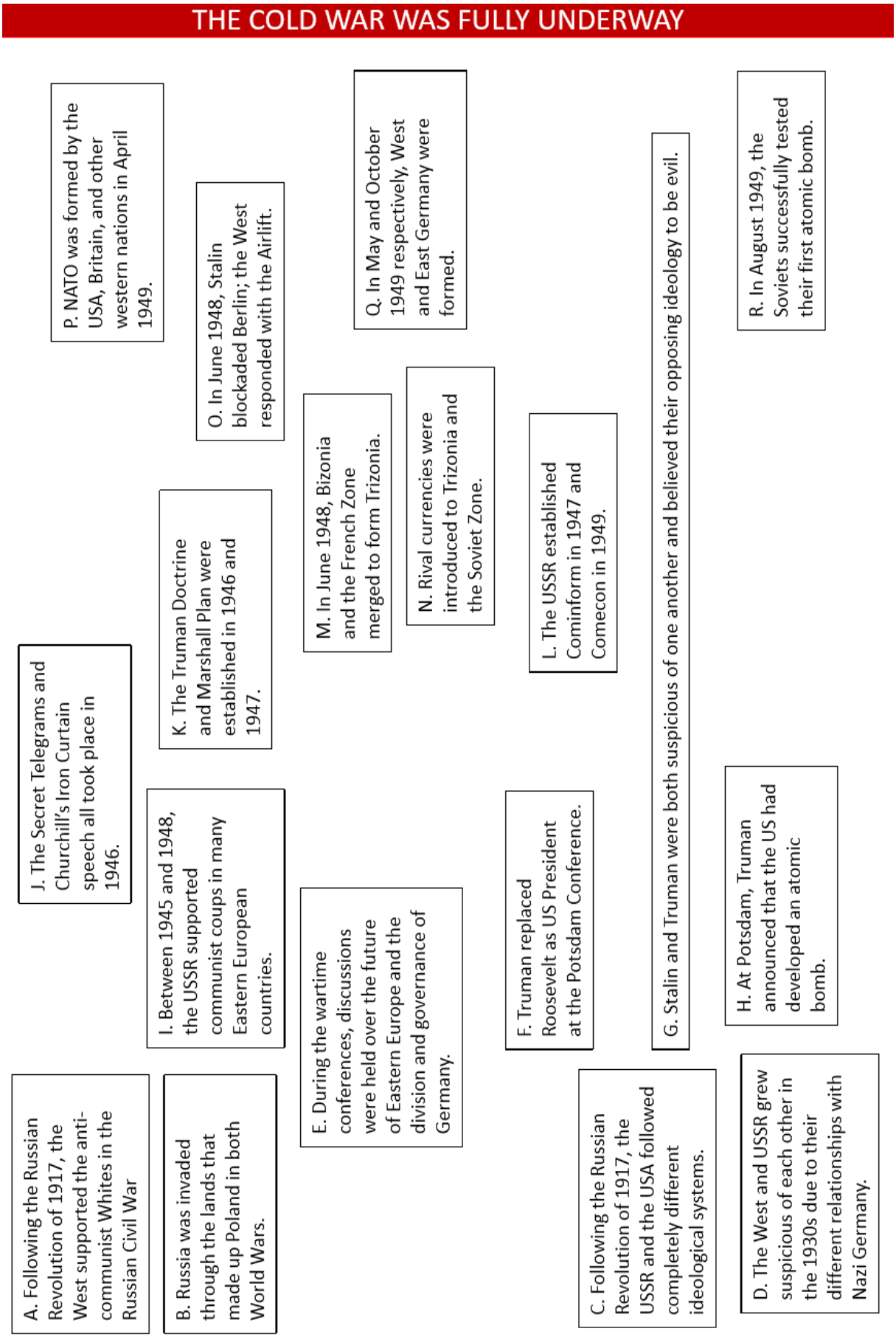
Sophisticated links between factors

Sophisticated links within factors

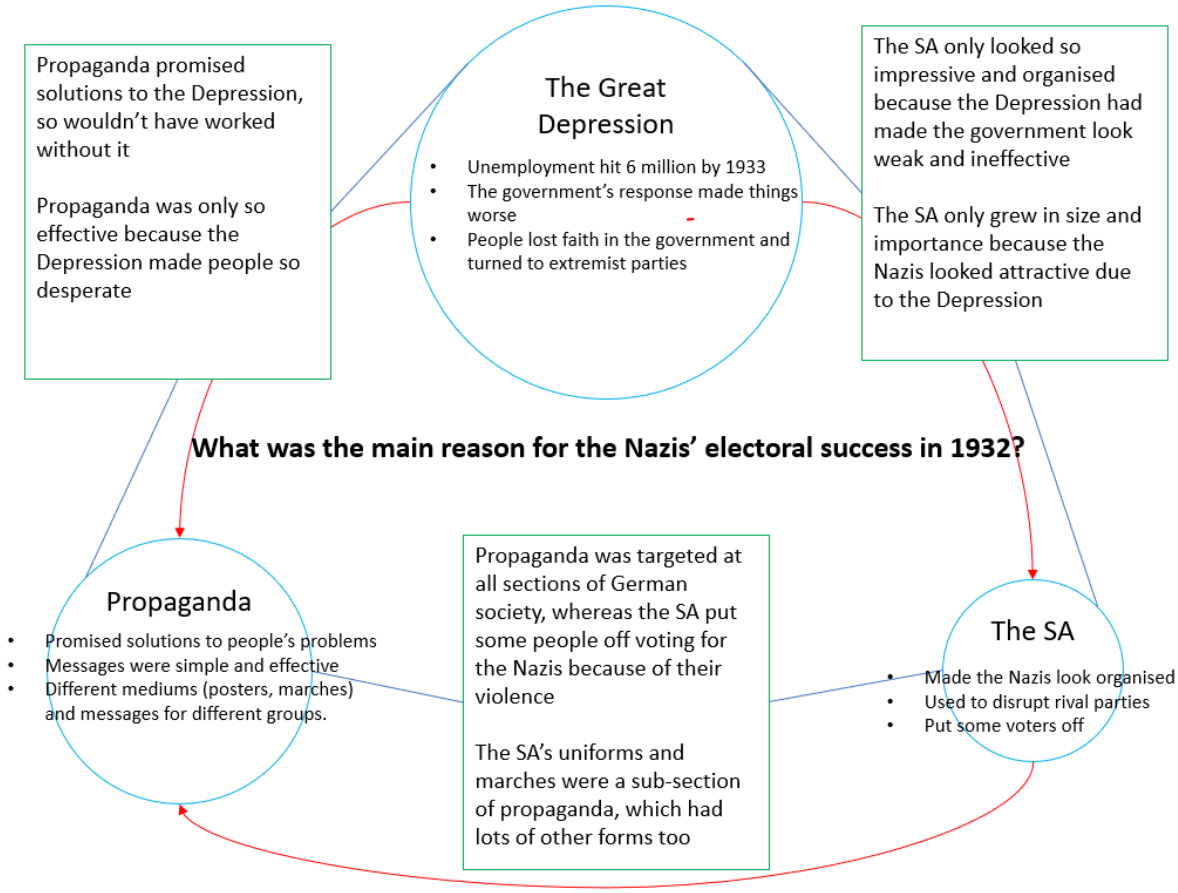
Links used to support a convincing argument

Appropriate causal language

APPENDIX F: Link Diagram Template: Development of the Cold War



Appendix G: Exemplar Circle Diagram from Intervention 2



The larger the circle, the more important the factor
Green boxes explain the comparative sizes of the circles
Red lines indicate a causal relationship between factors
Bullet-point lists contain key information about each factor, to be ticked during completion of the essay

Appendix H: Example of coding of work from Intervention 2

'The main reason for the development of the Cold War, 1945-1949, was disputes over Germany.'

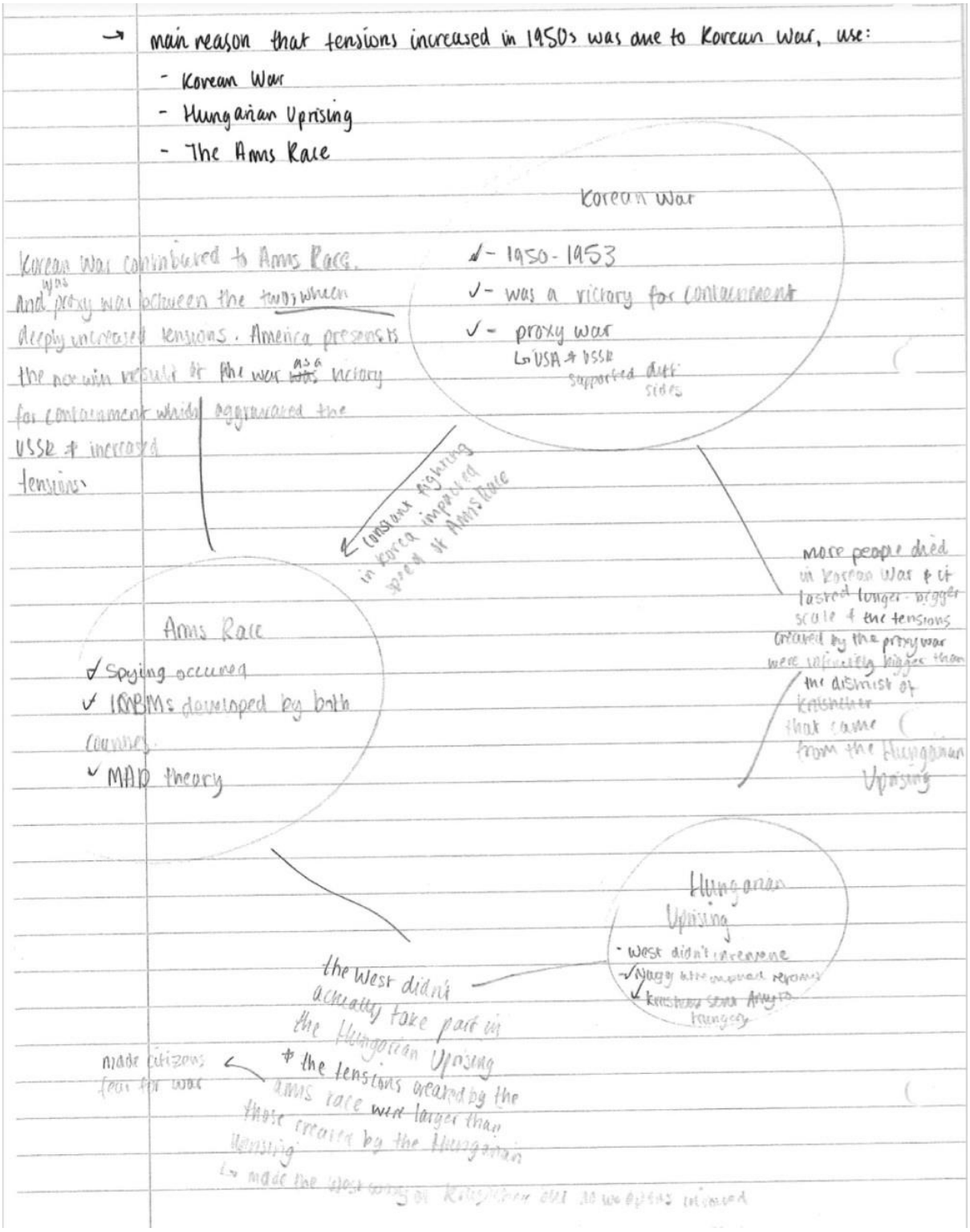
How far do you agree?

Between 1945 and 1949, there were many different factors which aided the development of the Cold War, including The Marshall Plan, disputes over Germany and the beliefs and personalities of Truman and Stalin. However, the beliefs and personalities of both leaders were the main reason for the development of the Cold War as it was a necessary cause, and without this, there would have been no consequences which caused the Cold War to advance.

The hugely differing beliefs and personalities between both the US leader, Henry Truman, and leader of the USSR, Joseph Stalin, played a significant role in the development of the Cold War, and were arguably the main reason that it began. Stalin believed in the spread of communist ideology across Eastern Europe, and was determined in his aim to establish a buffer zone between Germany and the USSR to prevent any future attacks from Germany which could potentially damage the Soviet Union. However, Truman held a drastically different view, believing instead in a democracy and the containment of communism. These contrasting views caused significant amounts of tension and suspicion between the two countries, and this suspicion played a monumental role in the events leading up to, and contributing to, the Cold War; for example, the formation of anti-communist movements such as The Truman Doctrine and NATO, both a result of Truman and Stalin's differing views. The outcome of both was dramatically heightened tension and suspicion, an increased desire for both countries to achieve their goals, and ultimately the development of the Cold War. As the personalities and beliefs of the leaders were the direct cause of these events, it is shown to be the most important factor, further confirmed through it being a necessary cause of the other competing factors, The Marshall Plan and disputes over Germany. Any disputes and disagreements over Germany, were a result of differing views, with Truman wanting free elections and a democracy for a defeated Germany, but Stalin trying to achieve a communist Germany, as well as having them pay reparations so high that the country would be crippled. Furthermore, their differing beliefs on the future of Germany led to division of Germany into zones and Berlin into sectors, and therefore the formation of Trizonia, the establishment of The Berlin Blockade, and finally the Cold War being fully underway. Moreover, this factor is significantly more important than The Marshall Plan, as without Truman wanting to prevent the spread of communism and strengthen the idea of democracy in Eastern Europe, as a result of his contrasting beliefs, there would have been no need for the formation of The Marshall Plan.

Sophisticated links between factors	Sophisticated links within factors	Links used to support a convincing argument	Appropriate causal language	Basic causal language

Appendix I: An intervention class student's Circle Diagram completed during Intervention 2, with self-regulatory ticking of content included in her essay



Appendix J: Resources provided to students in Intervention 3

"The main reason for the Nazis' electoral success in 1932 was the Great Depression." How far do you agree?

The Great Depression, Nazi propaganda, and the SA all contributed to the Nazis' electoral success. However, the Great Depression was the most significant reason.

The Great Depression had a catastrophic impact on the German economy, triggering the recall of American loans and a surge in unemployment, which hit five million in July 1932. Poverty levels rose and people became desperate; this was accelerated by the government's decision to slash benefits and increase taxes. The governments of Bruning and von Papen both failed to address the Depression, which reduced their support from the German people. As people became increasingly desperate, they turned to extremist parties – the Nazis and the Communists – in the hope of a solution. As the Depression worsened, the Nazis' share of the votes increased, rising from 12 in 1928, before the Wall Street Crash, to 107 in 1930 and then to 230 in July 1932. The Great Depression is the main reason for the Nazis' electoral success because it was a precondition for their rise in popularity, as the voting figures show. Had the Depression never happened, it is highly unlikely people would have turned to Hitler as they would not have been in a desperate situation. It was more important than the other factors because it enabled their success; without the economic crisis, both propaganda and the SA would have been ineffective.

However, Nazi propaganda was still an important reason why the Nazis did so well in the 1932 elections. The Nazis made vague promises, such as how they would fix the Depression, and used simple slogans such as "Work and Bread" to appeal to the millions of unemployed Germans. Propaganda was spread in many ways, including through posters, rallies, and newspapers, which

'Led' tells you very little about the nature of cause and effect. In this example essay, there are no 'led's. As you read through, highlight the words and phrases that have been used instead of 'led' and annotate them to explain what they mean.

allowed the Nazis to target a wide range of people, increasing their support. Hitler was portrayed as a superman who could save Germany; this contrasted with the weak government politicians and persuaded many Germans to vote for the Nazis, as they thought Hitler could solve all their problems. Nazi propaganda was an important contributory cause to their electoral success; it acted as a catalyst, determining the number of people who were voting for the Nazis in the hope of escaping the effects of the Depression. However, it was not as important as the Depression, since so much propaganda took advantage of the economic situation.

Finally, the work of the SA also contributed to the electoral success of the Nazis. The SA were the Nazis' private army and were used to disrupt the activities of rival political parties, especially the Communists, which made the Nazis a more effective and visible choice for voters. Furthermore, the SA were used for propaganda purposes, taking part in marches and rallies. Their uniforms and strict discipline presented an image of order to the people of Germany. This contrasted with the chaos of the Depression and encouraged people to vote for the Nazis. Despite this, the SA also negatively impacted the Nazis; some Germans, particularly the middle classes, were discouraged from voting for the Nazis because they disapproved of the violent behaviour of the SA. Whilst the SA did help the Nazis increase the number of votes, it is the least important reason behind their success. This is because, as with propaganda, the SA's effectiveness was determined by the Depression. Furthermore, whilst the SA contributed to the success of the Nazis' propaganda, they were only one small part of it, and the other forms would still have gained the Nazis votes anyway.

Whilst all three factors contributed to the Nazis' electoral success, the Great Depression was the most important because it created the underlying conditions which were necessary for both the SA and propaganda to persuade desperate voters to support the Nazis.

Word/phrase	Definition	Historical example from what we have studied
Trigger / triggered		
Catalyst		
Necessary cause		
Contributory cause		
Underlying condition		
Precondition		
Determined		
Enabled		
Accelerated		
Sufficient		
Latent		
Influenced		
Spark / sparked		
Exacerbated		
Tipping point		
Laid the foundations		
Turning point		

"The main reason for the unpopularity of the Weimar Republic, 1918-1923, was hyperinflation." How far do you agree?

Hyperinflation, the Treaty of Versailles, and the Weimar Constitution all made the Weimar Republic unpopular. The Treaty of Versailles was the main reason.

The Treaty of Versailles led to many problems for the Weimar Government. The Treaty forced Germany to pay the victorious powers £6.6 billion in reparations. This damaged the German economy and, when Germany failed to pay, led to the French occupation of the Ruhr, which led to hyperinflation. The Treaty also limited Germany's army to 100,000 men, which led to a massive drop in national pride, and a decrease in popular support for the Government. Finally, the Treaty led to people believing the Stab in the Back Myth, causing them to refer to the government as the November Criminals for signing such a harmful treaty. Ultimately, the Treaty of Versailles is the main reason for the unpopularity of the Weimar Republic. This is because it led to other issues, such as hyperinflation. Without the Treaty, it is likely that many of Germany's problems would not have happened, or if they had, that they would have been far less serious.

However, hyperinflation was also a major source of unpopularity for the Weimar Government. The French invasion of the Ruhr led to hyperinflation because the government printed more and more money to pay the striking workers. This led to money becoming worthless, which made life difficult for ordinary Germans. The middle classes and pensioners were affected particularly badly, as their savings became worthless. As it was the government that had caused this issue, people blamed them, and the Weimar Republic became even more unpopular. Although hyperinflation caused a great deal of unpopularity for the government, it was not the main reason. The government was already unpopular before hyperinflation set in; this

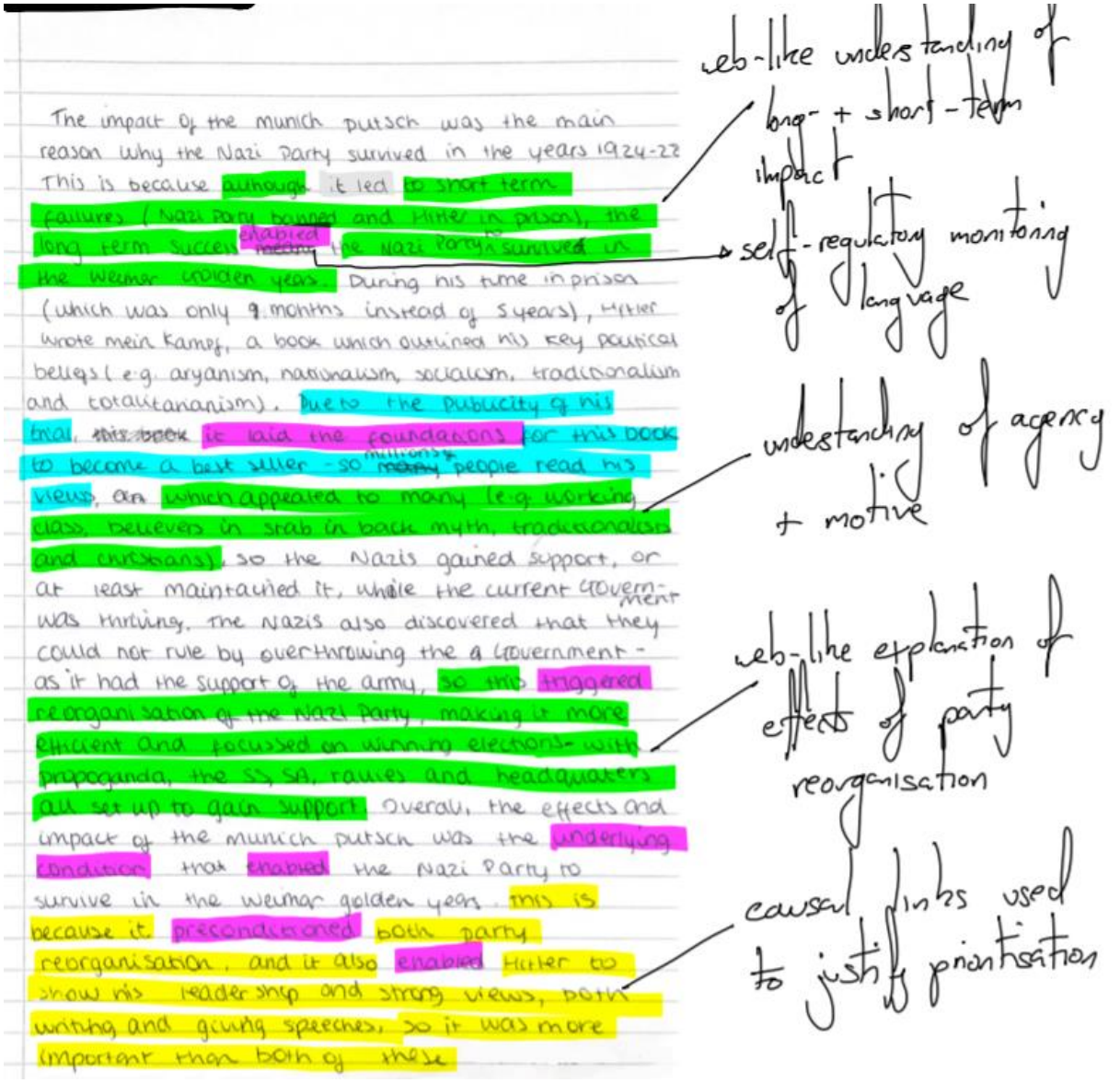
This is a decent essay, but its vocabulary could be much better. Using some of the examples from your table, annotate the paragraphs below to show where you would change the vocabulary and/or argument to better explain cause and effect. You can change individual words, or re-word entire sentences.

was because of the Treaty of Versailles, which also led to hyperinflation.

Finally, the Weimar Constitution also caused the government to be unpopular. The Constitution set out the rules for how Germany was to be governed, including that the voting system would be proportional representation. This caused weak, coalition governments to form, who often could not agree on solutions to Germany's many problems. This led many Germans to lose faith in the government, because it was unable to respond effectively to the problems caused by the Treaty of Versailles and hyperinflation. If the Constitution had produced stronger governments, then maybe the Weimar Government would have addressed the problems in a better way, causing it to be more popular. Although this is important, it is less important than the other two factors because it did not directly lead to unpopularity, but just made the other issues worse.

Overall, the Treaty of Versailles is the main reason that the Weimar Republic was unpopular because it led to another great source of unpopularity, hyperinflation, as well as causing many other problems for Germany.

Appendix K: Example of coding of work from end of project analysis



The impact of the Munich Putsch was the main reason why the Nazi Party survived in the years 1924-28. This is because although it led to short term failures (Nazi Party banned and Hitler in prison), the long term success enabled the Nazi Party to survive in the Weimar golden years. During his time in prison (which was only 9 months instead of 5 years), Hitler wrote Mein Kampf, a book which outlined his key political beliefs (e.g. aryanism, nationalism, socialism, traditionalism and totalitanianism). Due to the publicity of his trial, this book it laid the foundations for this book to become a best seller - so many millions people read his views, as which appealed to many (e.g. working class, believers in stab in back myth, traditionalists and christians), so the Nazis gained support, or at least maintained it, while the current Government was thriving. The Nazis also discovered that they could not rule by overthrowing the government - as it had the support of the army, so this triggered reorganisation of the Nazi Party, making it more efficient and focussed on winning elections - with propoganda, the SS, SA, rallies and headquaters all set up to gain support. Overall, the effects and impact of the Munich Putsch was the underlying condition that enabled the Nazi Party to survive in the Weimar golden years. This is because it preconditioned both party reorganisation, and it also enabled Hitler to show his leadership and strong views, both writing and giving speeches, so it was more important than both of these

Sophisticated links between factors	Sophisticated links within factors	Links used to support a convincing argument	Appropriate causal language	Basic causal language

Appendix L: Endpoint student survey

NB. Questions 17 to 21 were only issued to the intervention class.

Page 1

* 1. Writing arguments in History is easy

- Strongly agree
- Agree
- Disagree
- Strongly disagree

* 2. There is a "right" argument to questions in History

- Agree
- Disagree

* 3. I prefer to be told by my teacher what to argue when answering questions in History

- Strongly agree
- Agree
- Disagree
- Strongly disagree

* 4. I find it difficult to decide which factor is the most important when writing an essay in History

- Strongly agree
- Agree
- Disagree
- Strongly disagree

* 5. I often change my mind about my argument when writing in History

- Strongly agree
- Agree
- Disagree
- Strongly disagree

* 6. I feel like I have a range of tools and strategies to help me build my arguments in History

- Strongly agree
- Agree
- Disagree
- Strongly disagree

7. Please list any tools and strategies you have to help you build arguments in History

8. I have got better at constructing arguments in my History essays this year.

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

9. I have got better at constructing arguments **independently** in my History essays this year.

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

* 10. I am confident at answering questions that deal with causation (why something happened)

- Strongly agree
- Agree
- Disagree
- Strongly disagree

* 11. Causation questions require a different argument to non-causation questions

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

* 12. What kinds of justification could be used to argue for the most important reason something happened?

* 13. I am confident at explaining what causation means in History

- Strongly agree
- Agree
- Disagree
- Strongly disagree

14. My understanding of causation has improved this year.

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

15. What aspects of writing an essay with an argument (part c questions) do you find EASY in History? Give as many points as you are able to.

16. What aspects of writing an essay with an argument (part c questions) do you find DIFFICULT in History? Give as many points as you are able to.

17. I have found link diagrams (connecting boxes and explaining why) useful for improving my understanding of causation

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

18. I have found circle diagrams useful for thinking about which factor is most important

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

19. I have found circle diagrams useful as a planning tool

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

20. I have found vocabulary tables and example essays useful for explaining my ideas

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

21. Please give any feedback (positive or negative) you have on the link diagrams, circle diagrams, and vocabulary exercises.

Appendix M: Two-tailed t-test data

Scores: 2 | 1 | 0 | -1 | -2 — Sig-level = 5%

Two-tailed t-tests, using the scores per answer and significance level stated

Question	Baseline - intervention									Baseline - non-intervention									Difference between groups at Baseline			
	SA	A	N	D	SD	n	Mean	Var		SA	A	N	D	SD	n	Mean	Var		df	T statistic	sig?	p-value
1	1	12	0	4	0	17	0.588235	0.83044983		0	13	0	3	0	16	0.625	0.609375		30.742026	-0.12469	No	0.902
2	0	0	0	17	0	17	-1	0		0	0	0	16	0	16	-1	0		31	#DIV/0!	No	1.000
3	2	5	0	8	2	17	-0.17647	1.67474048		0	7	0	8	1	16	-0.1875	1.152344		30.536837	0.026708	No	0.979
4	0	11	0	5	1	17	0.235294	1.12110727		2	6	0	8	0	16	0.125	1.359375		30.241849	0.28392	No	0.778
5	1	9	0	5	2	17	0.117647	1.51557093		0	10	0	6	0	16	0.25	0.9375		30.082319	-0.34433	No	0.733
6	5	9	0	3	0	17	0.941176	0.99653979		2	14	0	0	0	16	1.125	0.109375		19.664008	-0.7185	No	0.481
10	4	10	0	2	1	17	0.823529	1.20415225		3	10	0	2	0	15	0.933333	0.728889		29.575359	-0.31774	No	0.753
11	0	8	8	0	0	16	0.5	0.25		3	4	7	1	0	15	0.6	0.773333		21.894901	-0.38581	No	0.704
13	4	6	0	4	3	17	0.235294	2.17993308		1	9	0	5	1	16	0.25	1.3125		29.946467	-0.03207	No	0.975

Question	End-point - intervention									End-point - non-intervention									Difference between groups at End-point			
	SA	A	N	D	SD	n	Mean	Var		SA	A	N	D	SD	n	Mean	Var		df	T statistic	sig?	p-value
1	1	7	0	9	0	17	0	1.17647059		2	10	0	2	0	14	0.857143	0.693878		28.888006	-2.48717	Yes	0.019
2	0	1	0	16	0	17	-0.88235	0.22145329		0	4	0	10	0	14	-0.42857	0.816327		18.699133	-1.699	No	0.107
3	4	8	0	5	0	17	0.647059	1.28719723		7	5	0	2	0	14	1.214286	1.02551		28.780365	-1.46964	No	0.153
4	2	9	0	5	1	17	0.352941	1.40484429		0	5	0	9	0	14	-0.28571	0.918367		28.99612	1.658786	No	0.108
5	1	10	0	5	1	17	0.294118	1.26643599		0	5	0	8	1	14	-0.35714	1.086735		28.556109	1.669788	No	0.106
6	5	8	0	4	0	17	0.823529	1.20415225		3	8	0	3	0	14	0.785714	1.02551		28.582352	0.099623	No	0.921
10	3	12	0	2	0	17	0.941176	0.64359862		1	14	0	0	0	15	1.066667	0.062222		19.431686	-0.61228	No	0.548
11	1	8	8	0	0	17	0.588235	0.35986159		0	6	8	0	0	14	0.428571	0.244898		28.997944	0.812026	No	0.424
13	2	13	0	2	0	17	0.882353	0.57439446		0	8	0	6	0	14	0.142857	0.979592		24.033094	2.295743	Yes	0.031
8	4	11	2	0	0	17	1.117647	0.33910035		7	7	0	0	0	14	1.5	0.25		28.932155	-1.9665	No	0.059
9	2	12	2	1	0	17	0.882353	0.4567474		3	8	3	0	0	14	1	0.428571		28.190029	-0.49071	No	0.627
14	2	15	0	0	0	17	1.117647	0.10380623		0	10	0	4	0	14	0.428571	0.816327		15.725225	2.715016	Yes	0.016

Question	Difference over time for intervention group				Difference over time for non-intervention group			
	df	T statistic	sig?	p-value	df	T statistic	sig?	p-value
1	31.07621	1.712026	No	0.097	26.893	-0.78412	No	0.440
2	16	-1.03078	No	0.318	13	-2.36643	Yes	0.034
3	31.4614	-1.97295	No	0.057	27.82152	-3.67782	Yes	0.001
4	31.60126	-0.30521	No	0.762	27.90828	1.058491	No	0.299
5	31.74541	-0.43623	No	0.666	26.79924	1.645031	No	0.112
6	31.71771	0.326984	No	0.746	15.42319	1.198907	No	0.249
10	29.30312	-0.35685	No	0.724	16.37295	-0.58058	No	0.570
11	30.57097	-0.46	No	0.649	22.34212	0.652389	No	0.521
13	23.88435	-1.60754	No	0.122	27.99824	0.274814	No	0.786

NB. To facilitate the calculations, the Likert scale was converted into a numerical one (+2 for strongly agree, +1 for agree, 0 for neither agree nor disagree, -1 for disagree, and -2 for strongly disagree).

The adjacent data presents a two-tailed t-test analysis for each question, with the null hypothesis being that the sample means of the compared groups are the same. As per convention, the significance value ("sig-level") is set at 5%.

The tests allow for comparison in four ways: intervention class baseline vs intervention class endpoint; intervention class baseline vs non-intervention class baseline; intervention class endpoint vs non-intervention class endpoint; and non-intervention class baseline vs non-intervention class endpoint.