Remembering More: Implementing classroom-based 'low stakes' retrieval activities to increase student recall of key knowledge in GCSE Geography

Rachael Speed

A Research & Development Project

Submitted for the MSc in Learning & Teaching 2022

DEPOSIT AND CONSULTATION OF THESIS

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

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

Surname	Speed
First Name	Rachael
Faculty Board	Education
Title of Dissertation	Remembering More: Implementing classroom-based 'low stakes' retrieval activities to increase student recall of key knowledge in GCSE Geography

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 Oxford Research Archives 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

Student success in examinations is at least partly dependent on their ability to recall large quantities of information during an exam. This study aimed to investigate whether the intervention of a low-stakes retrieval practice at the start of each lesson could be used to improve both student recall and confidence in applying knowledge to exam questions. Analysis of existing literature highlights strong positive effects of retrieval practice, but there is also a lack of research carried out by teachers in real classrooms using core-course material over extended periods of time. This study aims to provide additional insight. Student recall at the start of the study was assessed with a multiple-choice quiz, and students completed a questionnaire to elicit views on confidence and topics that needed immediate focus. Students completed three retrieval practice questions at the start of each lesson and were routinely re-tested with multiple-choice quizzes. Analysis of the data shows that both recall and confidence of the majority of the students increased over time. The study concludes that this data offers further support to the research literature by providing an example using core course materials and presents a workable intervention for use across the school in the next academic year.

Contents

1. Introduction

The reformed General Certificate of Secondary Education (GCSE) in Geography was first examined in 2018, and signalled a move to linear assessment, requiring that students sit all exams at the end of the course (Ofqual, 2018). Schools have the freedom to decide how and when they teach the content required by the exam specification, and in my school the curriculum dictates that all content for AQA Paper 1 (Living with the Physical Environment) is taught in Year 10, whilst Paper 2 (Challenges in the Human Environment) and Paper 3 (Geographical Applications) are taught in Year 11. This means that students are assessed at the end of two years on content they were taught over 18 months beforehand. This relatively long period, in combination with restricted curriculum time, means that teachers need to find some way to ensure that this content is kept fresh in students' minds to increase chances of success in examinations. The study explored here aims to explore whether the use of low-stakes retrieval practice at the start of each lesson can assist students in maintaining recall of Paper 1 knowledge throughout Year 11.

1.1 Remembering More

Learning is often identified with the acquisition or encoding of new knowledge (Karpicke, 2012), whilst the ability to recall knowledge is a means of assessing the knowledge each student has gained and retained (Karpicke and Grimaldi, 2012; Karpicke, 2012). The Ofsted Education Inspection Framework suggests that teaching should be designed to ensure that learners remember in the long term the content they have been taught (Ofsted, 2019), highlighting that learning is not just about

acquiring knowledge, but the ability to recall knowledge at a later date. Furthermore, student performance in end of GCSE exams is largely dependent on their ability to recall large amounts of information, whether that be factual, or skills based.

Retrieval practice is the name given to a method grounded in cognitive science that can be used to support students in strengthening their ability to recall information previously learned. Sometimes known as the testing effect, there are numerous definitions of retrieval practice in research literature, with the basic premise being that it is the act of recalling information to mind rather than rereading or hearing it (Roediger and Butler, 2011). Crucially, it is suggested that engaging in retrieval practice makes information stronger in the long-term memory, meaning that it should be easier to retrieve in the future (Bjork, 2012).

1.2 Wider Relevance

The issue of ensuring that students are able to recall knowledge and apply it in an exam situation is relevant to both my school and schools in general. Furthermore, the integration of approaches founded in cognitive science now forms a key part of the evidence that underpins the Ofsted Education Inspection Framework (Ofsted, 2019). However, evidence for the application of cognitive science principles, including retrieval practice, in the classroom is limited, and a detailed consideration of how these principles are used in lessons is critical to ensuring success (Perry *et al.,* 2021). Crucially for the study carried out here, UK-classroom-based studies from state-maintained schools where the study has been completed using core course materials are also lacking from the literature, as well as evidence from the implementation of

retrieval practice for GCSE Geography specifically. These issues will be further explored in the literature review section and suggest that this study can make some contribution to understanding of retrieval practice, particularly for designing a framework for implementation of a similar intervention in my school context.

1.3 Questions to Explore

Given the importance of applying cognitive science strategies to specific learning contexts (Perry *et al.,* 2021) the review of literature conducted in the next section needs to explore the methods that can be used to integrate retrieval practice effectively in the classroom. The literature review will also explore why retrieval practice works, and any negative aspects that exist that should be considered before wide-scale implementation. There also needs to be consideration of where the current evidence has come from and whether these findings can be deemed valid and reliable in the school context.

The following research questions are posed:

1. In what ways does retrieval practice influence Year 11 students' recall?
2. Does taking part in retrieval practice influence students' confidence in recall?

When reviewing the data collected in this study and determining the 'influence' that the intervention has had, there will be a number of factors to consider. The intervention itself will be explored further in the methodology chapter. The main limitations to consider are that the limited sample size (one class of 26 students) and lack of control group means that it will be difficult to judge the extent to which the retrieval practice intervention is responsible for changes in test performance.

2. Literature Review

2.1 Purpose and Scope

The purpose of this chapter is to conduct an in-depth review of the existing literature about retrieval practice in the classroom. An initial search was conducted in August 2021, whilst this study was in the proposal stages, using both SOLO (Search Oxford Libraries Online) and the Education Collection including ERIC databases. Table 1 shows the search criteria, filters and results yielded. Additional literature for review was then identified through citations of papers reviewed. Not all the studies identified in the initial search were read in full – on inspection many were found to only mention retrieval practice in passing, and therefore were not interrogated further.

Location	Search Terms	Filters	Number of results
SOLO	Title contains retrieval practice OR any field contains retriev OR any field contains testing effect	English language Teaching	136
Education Collection inc. ERIC	Anywhere contains retrieval practice OR anywhere contains retriev OR anywhere contains testing effect AND secondary school	United Kingdom	741

Table 1 Summary of literature search

Research into retrieval practice has been progressing rapidly since it was revived by Roediger and Karpicke in 2006 (Agarwal, Nunes and Blunt, 2021), so there is a wealth of literature to explore. My initial scrutiny of the literature available highlighted that studies have been carried out both in laboratory and classroom settings, at varied levels of education from primary school to university, and in many different subjects, although these are mostly UK and US-based. There are also several meta-analytic studies and literature reviews in existence. In order to ensure full coverage of the existing literature, I decided not to exclude studies that were laboratory-based or from

a different educational stage or subject than the one of interest to this study (secondary-level, GCSE geography). Instead, my reading and review of this literature should be critical, both of what the findings can contribute to this study and where its findings may not be applicable in this case.

My literature review will first look at cognitive science principles in the classroom more generally, before reviewing what retrieval practice is and why it works. I will then investigate the issues associated with retrieval practice in the classroom, before critiquing the current evidence base and its applicability to this study.

2.2 Cognitive Science in the Classroom

The Evidence and Practice Review from the Education Endowment Foundation (EEF) suggests that principles of cognitive science can have impacts on rates of students' learning (Perry *et al.,* 2021). However, they also suggest that evidence for the application of the principles derived from cognitive science in every day classrooms is limited (Perry *et al.,* 2021). It is important to acknowledge this gap, particularly as misinterpretation of the cognitive science research by serving practitioners could lead to poor implementation resulting in limited impact on learning (Alferink and Farmer-Dougan, 2010). Furthermore, given that the new Early Career Framework (ECF) places emphasis on the use of evidence-based practices in the classroom (DfE, 2019), which include cognitive science principles, it is necessary for individual practitioners to develop their own accounts of the implementation of cognitive science and its success in their particular subject, phase and learning context (Perry *et al.,* 2021).

In terms of education, there are two areas of cognitive science that have been especially influential. These are cognitive psychology, which is linked to research around mental processes, and cognitive neuroscience, study of the biological processes which underly cognition (Perry *et al.,* 2021). Ofsted identify practices that have their grounding in cognitive science; these include spaced practice, interleaving, retrieval practice, elaboration, and dual coding (Ofsted, 2019). Many of these ideas and practices are not new and have in fact been used in classrooms (albeit sometimes not explicitly named) for many years, however, there has been renewed interest in the contribution these practices can make to improved learning in recent years (Jones, 2019).

This review is not intended as a comprehensive assessment of possible ways of integrating cognitive science into the everyday classroom, but an exploration into the available evidence on one – retrieval practice, sometimes called the 'testing effect'. This method has been selected for this study because recent research has challenged the conventional view that tests simply measure knowledge, instead finding that tests produce a large positive outcome on learning (Karpicke, 2017), and particularly students' ability to retain information on a long timescale (Roediger and Karpicke, 2006), which is a fundamental goal of this study. Furthermore, testing is already widely used within classrooms (my own included) and integrating further quizzing into the classroom can be easily achieved (Roediger *et al.,* 2011). These factors all support the use of retrieval practice for this intervention, although the challenges will be explored later in this review.

2.3 What is retrieval practice?

The idea that testing has the potential to improve memory and retention is not a new one. In 1620, Bacon identified that it would be quicker to learn a text if you attempted to recite it and checked each time your memory failed, rather than just re-reading it (F. Bacon, 1620/2000, p.143, as cited in Roediger and Karpicke, 2006). The first study into the testing effect was published by Abbott in 1909 (Roediger *et al.,* 2011), in which the conclusion stated that 'the factor of recall is always an aid in the learning process' (Abbott, 1909, p.199). Since this study, the interest in retrieval practice and its utility for increasing retention has varied over time. Interest was revived with Roediger and Karpicke's (2006) experimental study into the impact of testing on memory (Agarwal, Nunes and Blunt, 2021) and it is now established through nearly 100 years of research. The overarching message from this research is that it is suggested that retrieval practice is a simple technique that can be easily implemented and lead to lasting positive results (Agarwal *et al.,* 2020). It is important to be aware that publication bias means that it is more likely studies with positive results have been reported, however.

There are many, broadly similar, definitions of retrieval practice in the research literature, with each study wording their working definition slightly differently. One of the most recent definitions, and one that has been derived from analysis of much of the extant literature comes from the EEF. They state that retrieval practice is 'the process of recalling information from memory with little or minimal prompting' (EEF, 2021, p. 21). This definition will be the one used throughout this study. There are many forms that retrieval practice can take in the classroom. Many of these have been

collated by Jones (2019; 2021a; 2021b), in a series of books published specifically for practicing classroom teachers. These methods include (but are by no means limited to) multiple-choice questions, short answer questions, problem solving, true/false questions, error spotting, labelling diagrams, image recognition and recitation of quotes (Jones, 2019; Perry *et al.,* 2021). Despite this, much of the research into the efficacy of retrieval practice has been derived from the use of multiple-choice and short answer quizzes only. The issues with the application and generalisability of the existing studies to everyday classroom practice will be explored later in this chapter.

2.4 Why does retrieval practice work?

The specific neuroscientific mechanisms underlying the success of retrieval practice remain unclear (Gao *et al.,* 2016), and it is important to note that the research emerging from cognitive psychology is often not situated within authentic educational contexts (Karpicke and Grimaldi, 2012). It is suggested that retrieval has an influence on learning in a number of ways. Firstly, when students take a test, the outcomes of this provides information on what they do (and do not) know which can guide future studying and teaching (Karpicke, 2017). Secondly, tests can influence students' motivation, as knowing about the test encourages students to increase study efforts, therefore leading to greater exposure to the materials (Roediger and Karpicke, 2006). However, these influences could be elicited by any number of classroom strategies, and it will seem obvious to educators that providing feedback will enhance learning (Karpicke, 2017). Therefore, it is the finding that taking the test in and of itself, without any prior or restudy or provision of feedback can improve learning that needs deeper theoretical explanation.

According to Atkinson and Shiffrin's (1968) multi-store model of memory, when material is learned it is transferred into the long-term memory. Theoretically, this is a store with endless capacity. However, for the knowledge to be used it must be retrieved from the long-term store and brought back into the working memory. One theory is that this retrieval of information leads to a strengthening of the memory trace, and creation of additional retrieval routes (Roediger and Butler, 2011), particularly if the retrieved knowledge is somehow integrated into new learning. The theory of 'desirable difficulties' (Bjork, 1994; Bjork and Bjork, 2011), suggests that it is the effort learners are required to make that increases the strength of memory – the degree to which knowledge is strengthened is proportional to the amount of effort involved during the retrieval attempt (Karpicke, 2017). Within the literature, there are many mechanisms proposed to account for the success of retrieval practice. A review of these will follow now.

2.4.1 Direct and Indirect Effects on Learning

Roediger and Karpicke (2006) suggest that retrieval has both direct and indirect effects on learning. Indirect (sometimes called mediated) effects, occur when it is not the act of taking the test itself that influences learning, but the fact that the testing has promoted learning via some other means, such as provision of feedback or motivation to study (Karpicke and Grimaldi, 2012; Roediger and Karpicke, 2006). Direct, or unmediated, effects occur when it is the testing itself that is producing learning.

In a review of evidence from laboratory tests conducted by Roediger and Karpicke (2006), the typical finding was that groups of participants who undergo repeated tests

following an initial test outperform groups whom either have no tests, or who are given time to re-study after the initial test. These results persist even when there is no feedback given after the initial testing phase (Roediger and Karpicke, 2006), and this therefore suggests that it is the completion of the test itself that leads to increase recall, rather than any indirect effect. Crucially, these findings have been replicated in more recent studies, with engagement in retrieval practice significantly improving performance on final tests compared to re-study or no-test control groups (see, for example, Agarwal, 2019; Roediger *et al.,* 2011).

In terms of comparison to restudy or repeated study, retrieval practice is well established in enhancing retention. This was demonstrated through Rowland's (2014) meta-analysis of retrieval practice research. Across the 159 studies reviewed, the overall effect size was $g = 0.50$ (representing medium effect) and 81% of comparisons favoured retrieval practice over the associated control conditions (Karpicke, 2017; Rowland, 2014). These results persisted regardless of the methods of retrieval used in the study.

2.4.2 Elaborative Retrieval Theory

The elaborative retrieval theory suggests that recalling a piece of knowledge (sometimes called a 'cue') can activate semantically related information that will further facilitate later retrieval (Gao *et al.,* 2016). This means that even information that is not actively recalled in a retrieval attempt is strengthened by its links to the knowledge that was recalled (Endres *et al.,* 2020). Carpenter (2009) suggests that this is because the act of recall activates memory traces, strengthening connections

between related concepts of that specific memory trace. It may also multiply the routes of retrieval for individual pieces of information (Roediger and Karpicke, 2006). The theory of elaborative retrieval is inconsistent with the cue overload principle, however (Watkins & Watkins, 1976). This suggests that the likelihood of successful retrieval decreases when the volume of linked knowledge to the cue increases (Gao *et al.,* 2016). A further criticism is that idea of elaboration is incompatible with retrieval-induced forgetting (Karpicke, 2017). This assumes that retrieval processes involve mechanisms that prohibit several additional items coming to mind. Furthermore, it is suggested that the retrieval effort can also place a memory into a state where it can be disrupted rather than enhanced. However, the understanding of the neuroscientific mechanisms behind this need further development (Roediger and Butler, 2011).

2.4.2.1 Retrieval Effort

The theory of elaborative retrieval is further complicated by the suggested interplay with the effort involved in the retrieval attempt. There is good evidence that conditions requiring more effortful retrieval confer greater benefits than easier conditions (Bjork and Bjork, 2011; Karpicke, 2017). Furthermore, more difficult retrieval also promotes better long-term learning (Agarwal *et al.,* 2020). However, the neuroscientific mechanism behind the benefits of retrieval effort is also unclear – it seems to work, but we do not yet know why.

2.4.3 Episodic Context Account

This theory, developed by Karpicke *et al.* (2014) suggests that retrieval practice improves later performance by enabling learners to develop more effective contextual

clues. This account suggests that whilst taking part in active retrieval, students try to mentally reinstate the study context and incorporate this with their current context. It is suggested that this then makes more distinct, stronger cues available for later retrieval (Gao *et al.,* 2016; Karpicke, 2017). Crucially, the episodic context account makes the explicit prediction that when retrieval is carried out intentionally, this leads to greater impacts on learning than incidental retrieval (Gao *et al.,* 2016; Karpicke and Zaromb, 2010). I suggest that this could have links to the retrieval effort theories. The more effortful the retrieval process, the more likely that students are thinking more explicitly about context of learning, leading to the strengthened memories suggested by this theory.

2.4.4 Transfer Appropriate Processing (TAP)

An older theory to account for the success of retrieval practice is transfer-appropriate processing. This suggests that performance on a memory task will be better when the processing required in initial learning matches that which will be required during a later final assessment (Morris *et al.,* 1977; Kolers and Roediger, 1984; Karpicke, 2017). It seems intuitive that practicing the skills required for assessment would lead to greater retention and therefore performance on later tests (Roediger and Karpicke, 2006; Karpicke, 2017), although this is yet another area whereby the cognitive architecture underpinning the effect is not well understood.

Conversely, there has also been investigation into situations where initial retrieval practice conditions are different to those carried out during final assessment (Karpicke, 2017). McDaniel *et al.* (2007) observed that retrieval practice conferred

benefits even where test formats did not exactly match the final test. Furthermore, a study by Gao *et al.* (2016) found that the largest effects of retrieval occurred following free recall or short-test formats, rather than after tests matching the final criterial assessment. This highlights further that the cognitive mechanisms behind the success of retrieval practice are not well understood, and current evidence can provide conflicting viewpoints as to why the strategy works to improve retention of previous learning.

In addition, transfer of retrieved learning to final test context can be difficult. Learners may not recognise that the knowledge they have acquired can be applied in a novel situation, or they may incorrectly apply the knowledge during transfer (Pan and Agarwal, 2020). Karpicke and Grimaldi's (2012) overview of research suggests that many students lack metacognitive awareness of the benefits of practicing retrieval. This is crucial for teachers to consider when deciding how to implement retrieval practice in the classroom. If research from cognitive psychology tells us that active retrieval can enhance learning, it is important to consider if students recognise this themselves (Karpicke and Grimaldi, 2012). Karpicke, Butler and Roediger (2009) found that students' judgements of success in learning are partially based on how easy they find it to process information. This can mean that students report feelings of confidence in learning using techniques such as re-reading and restudy, but this will not translate to performance in later tests (Karpicke and Grimaldi, 2012). Therefore, it is important that teachers fully explain and explicitly teach students how to use retrieval practice to be successful in learning, as metacognitive judgements have

consequences for student regulation and transfer of their learning (Karpicke and Grimaldi, 2012).

2.4.5 Theory of Disuse

Bjork and Bjork's (1992) theory of disuse assumes that representations of memory have two strengths – a storage strength, which reflects the permanence of a representation in memory; and retrieval strength which represents how available a memory is at a certain time (Karpicke, 2017). It is proposed that when an item with low retrieval strength is successfully retrieved, it increases in associated storage strength. This suggests that effortful retrieval of items with low retrieval strength improves learning by producing gains in storage strength (Bjork and Bjork, 1992). This provides further support that effortful retrieval conditions confer greater benefits than easier conditions (Karpicke, 2017).

However, this is only a redescription of retrieval practice, instead of an explanation of the reasons for it. In addition, the theory of disuse is grounded in psychological research, and it can be difficult to transfer relatively abstract theory to practice. This is further confounded by the fact that teachers are rarely experts in cognitive architecture and instead need to be informed how to correctly use retrieval practice techniques to foster deeper learning and understanding so that knowledge can be effectively strengthened, flexibly retrieved and transferred to novel situations (Roediger and Butler, 2011).

2.5 Issues with Retrieval Practice

This review has so far focused on the positive impacts that retrieval can have on long-term learning and retention. Before recommending the implementation of a strategy into real classrooms where students will be impacted, it is important ethically to explore the possible problems associated with the strategy and think about how these challenges could be mitigated against.

A criticism of retrieval practice is that it could lead to students experiencing increased test anxiety (Roediger *et al.,* 2011; Steele, 2011). This has important repercussions for the student's motivation and performance in a test and is an ethical dilemma – teachers are unlikely to want to engage in practices that could lead to students becoming uncomfortable in their lessons. Furthermore, it is likely that where students experience severe anxiety, this will lead to reduced positive or negative impacts on lasting learning (Khanna, 2015). The existing literature suggests that those students who do experience anxiety during testing benefit less from active retrieval practice than their peers who do not suffer (Karpicke, 2017).

It is fortunate, then, that numerous studies indicate that frequent classroom quizzing appears to reduce the level of test anxiety experienced by students (Leeming, 2002). Agarwal *et al.* (2014) surveyed 1408 middle and high school students in the US (roughly equivalent to secondary school age children in the UK). These students had all taken part in retrieval activities in their classrooms, and 92% of them reported that they believed retrieval practice had helped them to learn, whilst 72% said that taking part in regular quizzing helped them to feel less anxious about sitting tests (Agarwal *et*

al., 2014; Karpicke, 2017). The anxiety linked to retrieval practice quizzing in the classroom can be reduced by ensuring that the activities remain low stakes, meaning that the scores do not count towards any evaluation of the student (Agarwal *et al.,* 2014). This reduces associated anxiety because students do not feel under pressure to perform at all times.

A second criticism of retrieval practice is that it encourages rote learning rather than deep, meaningful understanding (Rowley and McCrudden, 2020). This is significant because rote learning is considered to be 'brittle and transient' (Karpicke and Grimaldi, 2012), and therefore students will be less likely to be able to transfer this learning to novel contexts in a test or exam scenario. The comparison of rote learning to meaningful learning can be linked to the similarity of the retrieval to the initial learning condition, as well as the timing of the first retrieval in relation to first learning. It is suggested that retrieval after only a very short delay is more representative of rote learning (Roediger and Butler, 2011). Despite these suggestions, there is evidence that knowledge embedded during retrieval practice is in fact flexible and can be transferred to new contexts, providing benefits to both high and low order thinking (Rowley and McCrudden, 2020).

Another criticism posited is that taking part in retrieval practice may result in students remembering the wrong answers. It is suggested that the use of multiple-choice questions as a vehicle for retrieval are particularly vulnerable to this, as they naturally introduce incorrect information that may then be assimilated into memory and retrieved in the long term (Karpicke, 2017). It is also necessary to ensure that multiple

choice questions are carefully constructed, so as to afford greater levels of effort in retrieval and promote learning (Karpicke, 2017; Marsh and Cantor, 2014). As a result of this suggestion, some authors suggest that short-answer questions will lead to greater learning as more effortful retrieval is required (Karpicke, 2017). However, in contradiction to this idea, Little *et al.* (2012) argued that multiple-choice questions can be designed to promote even greater retention than short answer questions. They posit that if questions are designed to have plausible 'lures' (incorrect answers), this will require students to think about why each answer is either correct or not (Little *et al.,* 2021; Marsh and Cantor, 2014). This will therefore have a positive impact on students' metacognitive awareness, as well as strengthening memory through mechanisms linked to retrieval effort.

A final factor to explore is when, and indeed if, to provide students with feedback following a retrieval attempt. Roediger and Karpicke's (2006) study reviewed laboratory evidence linked to testing and retention and found that when initial testing is followed by feedback this led to increased retention on final tests, even where these were delayed. Feedback is defined as a mediated effect of retrieval practice and is particularly useful in overcoming differences in individual student performance during initial retrieval (Karpicke, 2017). In particular, it is suggested that feedback is especially important in eliminating the negative effects of multiple-choice questions previously explored, as this allows teachers to pick up on where lures have been assimilated into memory (Butler and Roediger, 2008; Marsh and Cantor, 2014; Butler *et al.,* 2017). The timing of feedback may also have an impact. This is especially pertinent for classroom teachers to consider, as they will have to explicitly plan when to carry out

both retrieval and feedback activities within the available curriculum time. Marsh and Cantor (2014) suggest that delayed feedback comes with additional benefits, likely because it provides another opportunity for exposure to material. This undermines a conventional view in education that feedback must be given immediately to be effective (Roediger and Butler, 2011) and likely stems from the fact that delaying feedback acts as a spacing effect. Spacing is another cognitive science method which suggests pupil learning is improved when material is broken apart by intervals of time (EEF, 2021). This also therefore serves to highlight that it may not be possible to fully separate the effects delivered by different cognitive science approaches, as there may be complex interplays at work.

A further factor to consider is that the teacher must ensure that feedback is processed by the learner (Marsh and Cantor, 2014). This means that learners have to be actively engaged and listening, and it can be easy to lose focus when receiving feedback (Pan and Agarwal, 2020). This is significant both in ensuring that misconceptions or wrong answers do not persist and also in developing metacognitive awareness of the benefits of integrating retrieval (Agarwal *et al.,* 2020; Karpicke and Grimaldi, 2012). This is particularly pertinent given that researchers suggest students do not recognise the benefit of practicing active retrieval and do not use it as a study strategy outside of the classroom (Karpicke and Grimaldi, 2012).

This summary of issues with retrieval practice will be important to consider throughout the implementation of a retrieval practice intervention in this project and in further use of the strategy in the school.

2.6 Critique of Current Evidence Base

Although there has been considerable focus in recent years on extending the cognitive psychology research into retrieval practice into classroom contexts (Karpicke, 2017), there is still limited evidence for the application of this cognitive science principle in everyday classroom conditions (Karpicke and Grimaldi, 2012; Perry *et al.,* 2021). Additionally, much of the extant literature has been carried out in the USA, and is especially well established in laboratory-based studies which use undergraduates as participants (McDermott *et al.,* 2014; Rowley and McCrudden, 2020). There is a definite gap in the literature for studies (either laboratory or classroom based), that are based in primary and secondary (US K-12) settings, meaning that currently the application of basic scientific evidence may have low ecological validity, particularly for certain curriculum areas and pupil groups (Perry *et al.,* 2021). This will be important to address if educators are to effectively bridge the gap between cognitive science research and real educational practice (Rowley and McCrudden, 2020).

A second critique of the research currently in existence is that relatively few studies have been carried out in real educational settings. This is not to say that studies in classroom settings do not exist, but most of them have been carried out using material extraneous to the course (Roediger *et al.,* 2011). For example, an early study by Gates (1917) showed that testing improved retention of nonsense words, whilst Carpenter, Pashler and Vul (2006) tested students on their ability to recall word pairs. Roediger and Karpicke's (2006) review found that the majority of studies used word or picture lists (verbal learning tradition), which are rarely the methods used to assess students for national qualifications. This is a significant weakness when it comes to applying

research to the classroom. A major issue is that the students in these studies may feel that the material is arbitrary, whereas in real life, study is usually completed in pursuit of some larger goal (Kang and Pashler, 2014). This may mean that researchers are unable to account for the impacts that motivation has on efficacy of retrieval practice in the classroom. In particular, retrieval practice is often compared to some other scenario such as restudy. It is possible that passive strategies such as re-reading, which are usually used as an experimental control, are vulnerable to lack of attention caused by lower motivation (Szpunar, Khan and Schachter, 2013). Therefore, the success of retrieval reported in studies using passive controls may be artificially inflated if motivation is not taken into account (Kang and Pashler, 2014).

A further issue has been alluded to above – that retrieval practice is usually studied in relation to some passive learning strategy as a means of control. In many studies, the architecture of the experiment is the same. In an initial learning phase, students study a set of materials, and after this are allocated to one of a number of groups. In a retrieval condition, students will complete one or more activities designed to practice retrieval, whilst in a control condition, retrieval does not take place. Sometimes, more groups are included that vary the amount of retrieval, experience of feedback or final test taken (Karpicke, 2017; Roediger and Karpicke, 2006). This experimental model would be very difficult for teachers themselves to replicate. Firstly, it would take more time than is available, and perhaps more importantly, would come with serious ethical implications if the study used actual core course materials that students would be examined on at a later date. This also affects the ecological validity of applying the

findings of artificially constructed laboratory or classroom-based studies to a real educational context.

In addition, there are variations in the re-exposure participants will experience across conditions, and therefore, comparisons are biased in favour of repeated study conditions. Therefore, this creates data suggesting that there is no retrieval practice effect, or an advantage of repeated study (Karpicke, 2017). This would not be problematic if it were not in direct contradiction to the neuroscientific mechanisms of memory or was replicated in studies whereby delays in retesting also identify re-study as the most advantageous model (Karpicke, 2017), but this is not the case.

It is also important to remember that research on student learning and memory has a number of uncontrolled variables. These can include operational factors such as absence from lessons or study periods, external motivators or motivational style of the individuals involved, and individual differences in terms of cognitive architecture, for example working and long-term memory capacity (Agarwal *et al.,* 2017; Agarwal, Nunes and Blunt, 2021). Once again, this will have impacts on both internal and ecological validity, or the extent to which an intervention is implemented according to the intended procedure (Agarwal, Nunes and Blunt, 2021; O'Donnell, 2008). In particular, this is a factor that warrants consideration in the current study, given the relatively wide-ranging differences that are catered for in single classrooms in the school.

2.7 Implementing Retrieval Practice in the Classroom

2.7.1 Methods of Retrieval

Although retrieval practice has been widely studied, there is relatively little research about which types of techniques are best applied to lead to successful long-term retention (Endres *et al.,* 2020). A study by Rowland (2014) suggests that the use of different retrieval tasks makes little difference to the outcomes. Conversely, Endres *et al.* (2020) suggest that type of task does in fact make a difference in a study that compared short-answer to free-recall tasks. The authors suggest that desired educational goals should be considered when implementing a retrieval practice procedure. Problematically for serving teachers, this study links goals to the different theories of retrieval mechanisms, such as the elaborative retrieval theory and transfer appropriate processing (Endres *et al.,* 2020). I suggest that the majority of serving practitioners will be unaware of these theories, or at least have a very limited appreciation of them. Therefore, the educational goals of teachers are likely to be different to those of researchers, particularly where this study is concerned.

In terms of deciding the methods of retrieval to use in the classroom, it may be more prudent for educators to look outside of academic literature from the fields of psychology and cognitive science, and instead consult with the work of practicing teachers. Jones (2019; 2021a, 2021b) is a serving history teacher who has published a number of works that seek to make the cognitive science literature about retrieval practice accessible to teachers and provide them with a toolkit of resources and approaches that they can implement to engage students in retrieval. Often, one of the major factors that teachers will consider is how quick and easy the tasks are to

integrate into their lessons. In fact, it could be said that retrieval practice is already inherently present in good classroom practice – teachers pose questions for students to answer; they set tasks that require students to integrate past knowledge; and student sit exams requiring retrieval (Agarwal, Nunes and Blunt, 2021).

Despite this critique, there are some factors common in the research literature that teachers should be aware of. Firstly, is the need to keep the tasks designed for retrieval practice low- or no-stakes in order to reduce the negative impacts of test anxiety (Khanna, 2015). Secondly, is the need to provide feedback to students to overcome the issues of negative impacts of retrieval of wrong answers (Agarwal *et al.,* 2020) and thirdly is the need to ensure optimal task difficulty (Perry *et al.,* 2021) to ensure that the benefits of optimal retrieval are conferred.

The implementation of retrieval practice in the classroom should also be linked to subject-specific issues. As teachers, we are aware that each subject comes with differing knowledge and skills, and therefore mechanisms of retrieval practice may not be easily transferred between contexts. This is particularly important to consider in this study, given the relative lack of research literature existing for GCSE Geography in state-maintained schools specifically. Therefore, at all stages of the study it will be crucial to critically evaluate the wider applications of the evidence base, and continually review the methods used, in order to ensure that implementation is as impactful as possible.

2.7.2 Evaluating Success

The current literature that evaluates the success of retrieval practice does so based on studies that followed broadly similar experimental design. It is difficult, if not impossible, to replicate this in the classroom as a serving teacher, and therefore a different approach will be needed to assess the relative success of retrieval practice interventions for each individual group of students.

It seems logical that students should be assessed using some sort of test, given that the aim of retrieval practice is to improve performance in exams (at least it is in this study). The use of short-answer tests, where students are required to produce a short answer in response to question, is relatively unattractive to teachers because of the time it takes to administer and mark these (McDermott *et al.,* 2014). In comparison, multiple-choice question tests are more attractive because they are much easier and quicker to score (Karpicke, 2017). The potential pitfalls of the use of multiple-choice questions have been explored already. It is also not possible to separate the use of the multiple-choice test from its evaluative purpose from the inherent retrieval opportunity that is presented. Therefore, teachers should ensure that all suggestions from the current research are followed, namely ensuring optimal task difficulty, possibly through the use of plausible lures, and counteracting any negative effects of the provision of erroneous information through accurate and timely feedback to students (Butler and Roediger, 2008; Little *et al.,* 2012; Butler *et al.,* 2017; Karpicke, 2017; Perry *et al.,* 2021).

2.7.2.1 Impact of COVID-19

Retrieval practice is a strategy that seeks to elicit material from a student's long term memory and bring this into the working memory so that it can be used. This process relies on the information having been encoded in the first place. In the UK, school closures owing to the COVID-19 crisis resulted in significant disruption to education provision (Maldonado and De Witte, 2020) and therefore are likely to have had some impact on the initial encoding of information that teachers may now be expecting students to retrieve.

Studies into the impact of lost learning time as a result of COVID-19 are still in their infancy and draw heavily on existing data linked to other reasons for disrupted learning, such as summer learning loss and teacher strikes (Maldonado and De Witte, 2020). A major difference between these situations and the COVID-19 school closures are that there was widespread use of digital tools, such as video conferencing designed to ensure continuity of educational provision (EEF, 2020; Maldonado and De Witte, 2020). Despite this, initial evidence comparing test scores for students in primary schools from 2020 to those from 2019 documents substantial decreases in attainment (Blainey *et al.,* 2020).

In terms of evaluating the success of retrieval practice interventions, there are a few factors that should be considered here, especially given the context of the school used in this study. Firstly, is the fact that school closures are thought to have had a greater negative effect on students from disadvantaged backgrounds (Blainey *et al.,* 2020; EEF, 2020), possibly as a result of the impact of parental involvement (or lack of) (EEF,

2020) and also the impediment to academic, social and emotional learning (Storey and Zhang, 2021). Additionally, school attendance continues to be lower than the average reported in school census data from before 2019 (88.6% in March 2022, compared to a usual 95% (DfE Education Statistics cited in EEF, 2020). This is both a symptom of isolation due to COVID-19 infection, but also increases in anxiety, disengagement from education and other health needs (Ofsted, 2022). This will influence outcomes of any retrieval intervention, as students will miss opportunities for re-exposure to materials, testing, and feedback.

2.9 Summary

The review of literature presented here has served to highlight that this study can contribute to the gap in existing evidence from classroom teachers implementing retrieval practice activities based on core course material rather than arbitrary content. The review also highlights the importance of providing feedback alongside the retrieval activities, meaning that this will need careful consideration throughout. Furthermore, the extant literature has focused more on the implications of anxiety arising from classroom testing rather than confidence, so this study may also have something to contribute in relation to this factor. This study can also provide further commentary on the implications of individual differences such as Special Educational Needs and Disabilities (SEND), socioeconomic disadvantage, English as an additional language (EAL), absence as a result of COVID-19 induced effects, and loss of learning experienced as a result of national school closures in 2020 and 2021.

3. Methodology

3.1 Summary

This project was a small-scale practitioner enquiry, focused on one target class. There were two parts – firstly, implementing a programme of low-stakes retrieval activities, and secondly evaluating the impact that these have had on the recall of students, and their confidence in recall ability.

The project began with a multiple-choice test which covered a wide variety of the key knowledge the students had studied in Year 10. They then completed three retrieval questions each lesson, and were re-tested in December, February, and May.

As the project developed, changes were made to the original approach, including a change to the length of routine tests, and the number of questions asked of students at the beginning of lessons. The rationale for this will be explored further later in this chapter.

3.2 Participants

This project was carried out with one mixed ability Year 11 Geography class, which I have taught since September 2020 - the start of their GCSE studies. At the beginning of the project in October 2021, there were 26 students in the class with a range of target grades, from Grade 2 to Grade 7 (FFT20). Table 2 and Table 3 demonstrate the demographics of this group. The group was chosen as I am their only teacher, and therefore have the agency to implement the intervention as desired.

Characteristics	Total in class	Percentage
Boys	16	61.5
Girls	10	38.4
EAL	5	19.2
SEN (K)	6	23.0
SEN (EHCP)	2	7.6
Pupil Premium (PP)	9	34.6

Table 2 Class demographics

Target Grade (FFT 20)	Number
1	0
2	2
3	7
4	4
5	5
6	6
7	2
8	0
9	0

Table 3 Target grade distribution

Analysis of the demographics of the study class identifies that only 5 students are not in a target group (EAL, SEND or PP). This is equivalent to less than 20%, meaning a large majority of the class will be seen as having additional needs that need to be targeted in teaching. Furthermore, almost 70% of the class has a target of Grade 5, or below. This means that recall of key facts such as key word definitions are particularly important for these students' performance in the final exams. The aim is that engagement in retrieval practice will contribute positively to this ability. During data analysis and discussion, reference to the relative performance of each of these groups will be made.

The intervention used involves a change in teaching approach, so all students were involved. Each of them took part in the lesson-by-lesson retrieval activities and completed pre- and post-test quizzes and questionnaires. This intervention will be further explored in the next section.

3.3 Intervention

Table 4 provides a summary of the activities that took place during the completion of the intervention and data collection.

Month	Intervention Activities	Data Collection	Changes made from original plan
October	Three questions at the start of each lesson on prior knowledge	Baseline quiz and questionnaire given (October 4th, 2021)	
November			
December		Multiple choice knowledge quiz (December 15th, 2021)	Quiz made shorter – five questions on each topic in the Paper 1 exam
January			Questions changed – one from Section A, one from Section B, one from Section C
February	Six questions at the start of each lesson on prior knowledge	Multiple choice knowledge quiz (February 28th, 2022 – *delayed from February 18th, 2022, due to Storm Eunice school closure*)	Number of questions increased to six – one from each section of the Paper 1 and 2 exams and a skills-based question
March			Printing the questions off
April		Planned quiz not carried out.	
May		Final quiz and questionnaire	10 questions for each topic in Paper 1 (60 questions in total).

Table 4 Summary of research schedule

The intervention involved students completing three prior knowledge questions at the start of every lesson (three times a week). This intervention was chosen from the many possible ways of implementing retrieval practice explored in the literature review because of the ease of integrating it into everyday classroom practice. Furthermore, studies have shown that taking part in short-answer quizzes on a regular basis with immediate feedback is an effective way to ensure benefits of retrieval

practice are made available to all students (Gurung and Burns, 2018; McDermott et al., 2014).

Students in the class are well adapted to the start of lesson routine in my class. They come in and get out their books, and then write and underline the date and title which is always displayed on the board. Given this, it was relatively simple for me to add a stage to this, whereby they then answer the retrieval questions, which I also displayed on the first slide. An example of this is shown in Figure 1.

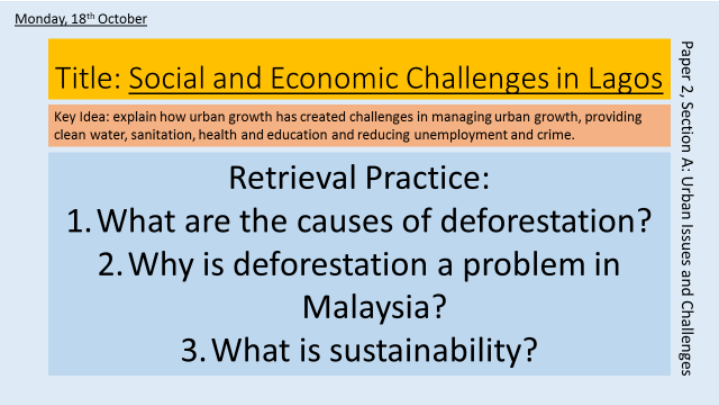

Figure 1 Example retrieval questions (October)

Students were told that if they did not know the answer, they should write out the question instead. This was intended to reduce the opportunity for students to 'opt out' of completing the tasks, as they would have to write something regardless of if they knew the answer. They were then encouraged to write the answer next to the question in a different coloured pen when answers were gone through in the class. This then allowed me to have a quick visual indication of poorly answered questions, which can then be repeated in subsequent lessons to ensure that students are then retrieving the correct answers. However, I made it clear that there were no judgements to students for not having got the answers correct, and I did not ask students to say how many answers they had achieved at any point. This was important to maintaining the 'low-stakes' aspect of the intervention, which is important to

preventing an increase in test-anxiety (Karpicke, 2017). In addition, a low or no-stakes approach to retrieval practice enables students to see this process as useful for learning, rather than a test of retained knowledge (Agarwal, 2019).

To increase the mnemonic benefits of testing (Roediger and Butler, 2011), the correct answers to questions were shared immediately following the retrieval attempt. Usually, I asked students to share their answers, therefore giving me an opportunity to address misconceptions or incorrect retrieval. Timely feedback, whether given after each question or at end of the retrieval practice phase, is identified in the literature as key to ensuring positive effect of testing and can also counteract any negative effects that stem from an incorrect retrieval in the first place (Roediger and Butler, 2011; Marsh and Cantor, 2014). Students were encouraged to record correct answers to any questions they had either not answered or answered incorrectly so that notes could be referred to later.

3.3.1 Deciding the Questions to Ask

Initially, I decided which questions to ask based on performance in the original knowledge test at the start of the intervention. When marking these tests, I recorded whether each individual had got each question correct or not. The design of the test, which split questions up by topic, therefore allowed me to see which topic was most poorly answered. I initially focused, therefore, on content from the Hot Deserts and River Landscapes topics, as these were the topics with the lowest scores.

However, I later decided that it would be more beneficial to students if they were asked questions on all topics with increased regularity. Ethically, it was important to ensure that students had exposure to all topics in Paper 1 to increase their chance of performing well in their final GCSE examination. The changes related to this decision are discussed in the next section.

3.3.2 Changes to the Intervention

As the project continued, I decided to make changes to the intervention, partly as a result of collaboration with colleagues that will be further explored later in this chapter.

Firstly, it was decided that to ensure more systematic coverage of the Paper 1 content, I would ask one question from each of the sections at the start of every lesson (Section A: Natural Hazards, Section B: Living World, Section C: Physical Landscapes – Rivers and Coasts). These were always asked in this order (see Figure 2), so as to familiarise students with the layout of the paper and content in each section. This is linked to research on the transfer-appropriate processing effect, whereby retrieval offers greater gains if the activities completed during learning phases are similar to those in the testing phase (Roediger, 1990). In order to make greater use of

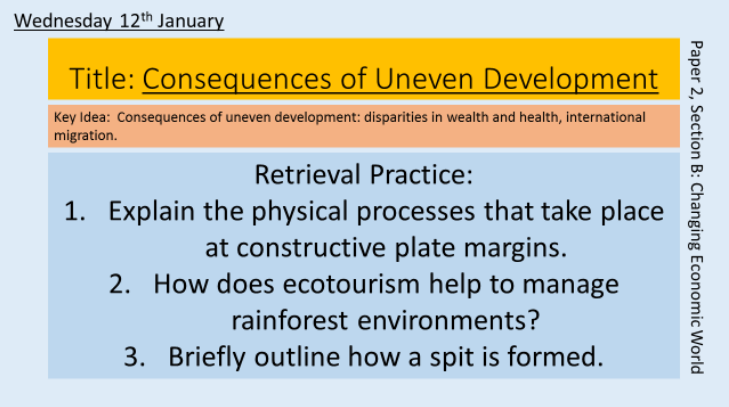

Figure 2 Example retrieval questions (January)

the transfer-appropriate processing effect, I also later started to use exam command words in the wording of questions to increase familiarity with this language.

In February, I finished teaching the class all the content needed for Paper 2 of the exam. At this point, I decided to adapt the retrieval practice to include content that students will need to answer in the Paper 2 exam (Section A: Urban Issues and Challenges, Section B: Changing Economic World) and skills questions that could appear in any of the three papers (see Figure 3). Again, this was an ethical decision to ensure that students had every opportunity to revise all content needed for their GCSE examinations.

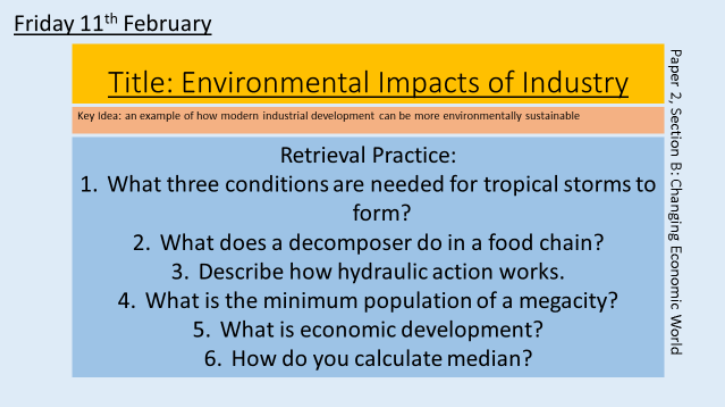

Figure 3 Example retrieval questions (February)

A final change made in consultation with students and staff was to provide students with a printed copy of the retrieval questions to stick into their books and then write answers next to. This occurred after the introduction of three more questions, as it was becoming more time consuming for students to write out the questions they didn't know.

3.4 Data Collection Methods

3.4.1 Pre-Intervention Quiz

At the beginning of the study, on October 4th, 2021, students were given an extensive 108 question multiple-choice test to gauge recall of knowledge from a wide range of the specification knowledge for Paper 1 (Appendix 1). Following review of the existing literature on retrieval practice, I decided that a multiple- choice quiz was the best way to assess initial performance. This is because it is a widely used method of testing student recall, partially due to the focus on the education system of the USA in research, where it is important to be able to grade large volumes of work in a short time period. This was also useful to me – the more quickly I could mark the initial quizzes and assess gaps in recall, the more quickly I could start to design retrieval practice questions to focus on the topics most in need of review.

3.4.2 Subsequent Knowledge Quizzes

In December, the students took a second quiz to determine if there had been a change in recall. Initially, it was planned that this would be similar to the original test, but I decided to shorten this to a 30-question quiz, with 10 questions for each of the exam sections A-C., for several reasons. Firstly, it took a long time for even the highest ability students to complete the original test, and maintaining student interest, effort and behaviour for this length of time was difficult, potentially affecting the scores gained. Secondly, I decided to focus on questions that had been asked in retrieval practice in the preceding weeks to instil confidence and build motivation. It was hoped that students would see the benefit of taking part in retrieval practice for their recall and this would engage those that had been less willing beforehand. I have not

encountered an approach like this in the research literature, however, given that most studies have not been carried out by serving practitioners using core-course material (Roediger *et al.,* 2011), this is not necessarily surprising. Instead, I have designed this intervention using my professional judgement, informed by the approaches taken in existing studies, with the appreciation that these cannot be replicated at the scale to which I am working.

A similar shorter test was repeated in February. Finally, a longer test consisting of 20 questions per topic was completed in May 2022, just before students sat the exam for Paper 1. Each of these testing points can be used to track relative performance of students over time and will be explored in the findings section.

3.4.3 Pre- and Post-Intervention Questionnaires

Two questionnaires were completed by the participants, one in October 2021 at the beginning of the study, and one in May 2022 at the end (Appendix 2).

The questionnaires consisted of mainly closed questions, using yes/no or Likert scale responses and were designed to elicit student views about their confidence in remembering and using knowledge from the prior year's teaching. I also asked some questions about revision strategies so that I knew the extent to which students would need coaching about more impactful methods.

3.4.4 Interviews

I initially planned to interview a cross-section of students from the class at the end of the study. This would have included males and females, SEND students and Pupil Premium students. However, due to changes made to the Year 11 school schedule with relatively late notice, this was later decided not to be possible. Instead, I asked students if they would be willing to volunteer to stay at lunchtime and answer some questions. Two students agreed to do this, and were interviewed separately, as per student request. This took the form of an unstructured discussion, in which the students were simply asked to share the views on retrieval practice for their confidence in recall and brief notes recorded by me.

3.4.5 Observation of Student Behaviour

As a teacher, I have the opportunity to observe students from the front of the room. This has allowed me to see the way in which students engaged with tasks on a day-to-day basis. Students in the class were also vocal on their opinions towards the tasks they were carrying out, giving me insight into their thoughts and feelings on the utility of retrieval practice tasks for their confidence and level of recall. Throughout the period of the study, I made notes about any individual behaviour that may be influencing scores in the quizzes, and also where day-to-day retrieval appeared to be having a large positive impact.

3.5 Analysis

3.5.1 Multiple Choice Quizzes

Multiple-choice quizzes will be marked by me as soon as possible after completion. Student responses to each question, whether correct (Y) or incorrect (N), were recorded in an Excel spreadsheet, and the total score and percentage noted. An example of this can be seen in Appendix 3.

3.5.1.1 Question-level Records

The question level records were used to determine areas that students need work on to improve recall. These questions were then the focus for initial retrieval practice at the beginning of the intervention. This also highlighted where students would benefit from further direct re-teaching of content, which could be included in curriculum planning outside of this intervention.

3.5.1.2 Measures of Central Tendency

Given the small sample size and simple percentage score data collected, the data has been analysed with measures of central tendency. I compared the mean average scores of the whole class across the four testing periods. This approach is similar to that observed in the literature, whereby students are assessed prior to and after completing retrieval activities (see, for example, Roediger and Karpicke, 2006; Karpicke, 2017). In this study I also compared the percentage change each individual student achieves from October 2021 to May 2022. This is something not seen in the research literature reviewed, possibly because these studies have used much larger sample sizes, and also because they have not been carried out (in most cases) by the

class teacher, so the researcher was unlikely to have knowledge of each participant as an individual. However, comparisons of performance for each individual overtime will add value to the discussion, particularly in seeking to understand patterns of success in improvement in recall. I will use qualitative data such as that elicited through observation of students in attempts to provide reasons for the patterns observed.

Additionally, the data elicited from multiple-choice quizzes will be visualised graphically to assist in the drawing of conclusions from the data.

3.5.3 Questionnaire Data

Closed questions were analysed in terms of number of responses for each category. Where the same questions are asked across both questionnaires the change in number of responses is assessed. Open questions linked to students' revision techniques and areas of weakness were first coded to different categories and then the number of responses to each category recorded. An example extracted from the original data analysis is shown in Table 5 below.

Is there any content from Year 10 that you feel is going to be particularly hard for you to remember?	Blank	Deserts	Coasts	Rivers	Hazards	Rainforests
	3	7	5	11	4	3

Table 5 Example of coding of questionnaire data

3.5.4 Interviews

Given the changes that occurred to the plan for interviews, these elicited relatively little detail. The comments made by the students will be used to provide support or challenge to the quantitative data analysis throughout the findings section.

3.6 Ethical Considerations

In order to ensure that this study was carried out with clear ethical intentions, I consulted the Ethical Guidelines for Educational Research, set out by the British Educational Research Association (BERA, 2018). An application including the title, research questions and methods of investigation was submitted to the Central University Research Ethics Committee (CUREC) in September 2021. In order to ensure full compliance with the *modus operandi* for the MLT course, a letter was submitted to my headteacher asking for their permission to complete the research (Appendix 4). In January 2022, this headteacher was replaced with a new one. I re-submitted the letter for their information, stating the aims of the study and what had been achieved so far. When all appropriate ethical approval had been granted, I then spoke to the class involved about the study. They were informed that they would be answering three retrieval practice questions at the start of the lesson, and that this would now be normal lesson routine. I also explained that they would be completing routine tests and completing questionnaires to contribute to my evaluation of the success of this intervention. Students were told that although they would be required to complete the tests, they could ask for their data to be removed from the analysis at any point, by simply approaching myself or another trusted member of staff. In this way, it was made clear that involvement in the analysis of the intervention was optional.

3.7 Collaboration

Collaboration occurred throughout and will continue to take place following the end of the project.

3.7.1 Learning Support Assistants (LSAs)

I was fortunate to have the support of the same LSA in all lessons during the study period. This LSA is in the room to support the two EHCP students but was also invaluable in discussion with me about changes to make to support the class as a whole. Following discussion with this staff member, I decided to add questions about Paper 2 once the teaching for this had finished. The LSA also suggested that I print the questions to stick in books to provide a permanent record of questions asked to support students likely to engage in independent work. The LSA also worked with me to observe the retrieval behaviour of certain students which will be explored later.

3.7.2 Teaching Staff

The findings of this study will be disseminated to all teaching and learning support staff through a session delivered as part of the schools' CPD Programme. The intention is that I will outline the intervention I have undertaken here and explore the findings. Staff will then be encouraged to adapt the intervention so that it is functional in their own classroom context, and review will take place as part of CPD to determine if any further changes can be made.

3.8 Limitations

The limitations of the study largely stem from the small sample size. Only one class, consisting of 26 students is involved in the study. This means that results will not be generalisable on any scale, whether to the school or to wider contexts. Furthermore, the class will not be compared to a control group, meaning that it will be impossible to determine whether my intervention is responsible for changes in recall to any extent.

Furthermore, within the study group, it will not be possible for me to separate the impact of the retrieval practice intervention at the start of lessons from any impacts of day-to-day exposure to material during teaching, homework or independent revision carried out by students. It would be entirely unethical (and largely impossible) to restrict opportunities for revision, re-study, or retrieval of Paper 1 content so as to ensure that the quizzes are only testing the impact of the retrieval practice intervention.

4. Findings and Discussion

The following section will present the findings of this study with reference to the relevant literature already reviewed. Firstly, I will review the results of initial multiple-choice study and review how this guided the intervention. I will then explore data related to the two research questions used to frame this study, starting with the impact of the intervention on recall and then moving on to look at confidence.

The main findings are as follows:

- 66% of students show an improvement in scores from the October 2021 quiz to May 2022 quiz.
- The mean average score for the whole class increased from 51.2% in October to 59.9% in May.
- Students reported an increase in confidence. In October, the majority (54.2%) of students reported their confidence in recall as 2. By May, the majority (45.8%) reported confidence at 4.
- 100% of students reported that they had found taking part in retrieval practice useful in improving their confidence in recalling content from Year 10.

4.1 Results from October Multiple-Choice Quiz

Students completed a 108-question multiple choice quiz on October 4th, 2021, in order to assess the level of recall they each had on the content taught in Year 10. This had two purposes – firstly, to function as a benchmark against which future performance could be assessed to determine success of intervention of retrieval practice, and

secondly to highlight where there were large gaps in recall so that this knowledge could be targeted with retrieval practice.

When carrying out the question-level analysis on the first quiz, I flagged any question which more than 10 students had got incorrect. This consists of over a third of the class and I judged this to be a good benchmark, as it encompasses a significant number of the students, without being too onerous to address within the intervention. However, review of the data revealed that 62 of the 108 questions completed fell into this category. On reflection, this is perhaps unsurprising given that the mean score for the whole class on this quiz was 51.2%. This highlights that the recall of key knowledge needed for the examination for Paper 1 in May 2022, was at a relatively low point at the beginning of the year, supporting the need for an intervention to increase recall.

In order to determine the topics to focus on at the beginning of the intervention, I analysed the number of questions answered incorrectly in each topic, the results of which are shown in Table 6.

Section of Paper 1	**Topic**	**Questions Answered Incorrectly**
A	Tectonic Hazards	4/19
	Nepal Case Study	3/5
	Chile Case Study	4/5
	Reducing Hazard Risk	1/4
	Weather and Climate	5/6
	Typhoon Haiyan Case Study	3/8
	Reducing Climate Change	4/4
B	Ecosystems	1/4
	Tropical Rainforest	3/8
	Hot Deserts	7/7

C	Rivers	14/16
	Coasts	10/20

Table 6 Incorrect answers by topic

This analysis shows that a number of topics were incorrectly answered across a number of students, so it is difficult to determine where initial focus should be made. However, two topics, Reducing Climate Change and Hot Deserts, stand out as having had **all** of their respective questions answered incorrectly by more than a third of the class. This suggests that these topics will require additional focus in retrieval practice in order to boost long term retention of this knowledge. However, I also posit that there is an additional factor at play here – the loss of learning as a result of COVID-19 school closures, which I will now explore.

Figure 4 shows a summary of the national closures and how this corresponds with school-based closures and the material being taught at each point.

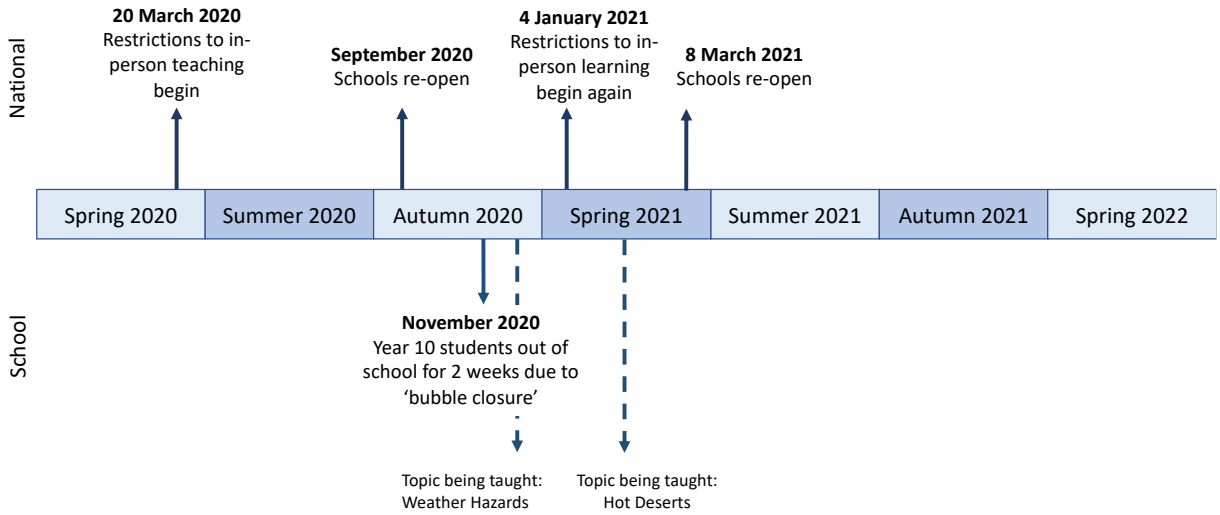

Figure 4 Timeline of COVID disruption

The timeline highlights that the two topics with the most incorrect answers in the quiz were those taught whilst students were taking part in remote learning. Anecdotally,

the lessons delivered via Google Meet were not well attended and the submission of work lacking from a number of the students, although I do not have access to records of attendance to quantitatively confirm this. Therefore, I was relatively unsurprised that so many questions were answered incorrectly in the Weather Hazards (which Reducing Climate Change is part of) and Hot Deserts topics. However, given that many students may not have engaged in initial learning, this is not an issue with recall, but rather original knowledge acquisition. The literature review carried out shows that retrieval practice has been studied for its ability to strengthen memory traces, not as a method for teaching knowledge. Therefore, missed learning opportunities need to be considered in the planning of retrieval questions, as many students will not have the knowledge stored in their long-term memory to retrieve. Instead, this material will need to be re-taught either in lesson time or with homework activities, before it becomes part of routine retrieval practice.

There are other factors which could be responsible for the large number of questions answered incorrectly by over a third of the class in the other topics. Firstly, the quiz was not carried out until October, and teaching of the content finished in July. Although not explored here, there is some evidence that students recall is reduced after the long summer break (something termed the 'summer learning loss effect' (Maldonado and De Witte, 2020)), as the knowledge they had previously stored in long-term memory has not undergone a cycle of recall and therefore the strength of the memory trace is diminished. Ideas about the strength of memory trace are linked to both the theory of disuse and elaborative retrieval effect that have been explored in the literature review.

A further factor resulting in incorrect answers could be a lack of engagement with the initial teaching in the classroom, and individual differences in student ability to encode information in their long-term memory and then to actively recall it. For example, the class that participated in the study had a total of seven students with SEND. Some evidence suggests that students with SEND diagnoses may struggle to hold as much information in their working memory as those without SEND, therefore limiting rate of transfer to long-term memory (De Weerdt, Desoete and Roeyers, 2013), and the volume of information that can be recalled. Again, this is beyond the scope of this study but may warrant further exploration. Furthermore, the initial test that students sat was very long. It is possible, therefore, that student engagement and motivation to complete this waned at certain points. This may account for the relatively high rate of incorrectly answered questions in the topics which were always at the end of the quiz. However, both of these suggestions are conjecture at this point. The sample size of both SEND and non-SEND students is simply too small to be able to draw reliable conclusions as to the impact of SEND on performance of students. Additionally, there is no available comparison of a shorter initial test to determine if the length of the test at the start of the intervention had some impact on performance at this stage.

Further light can be shed on the performance in the initial quiz by comparison with the areas of focus that the students identified in the pre-test questionnaire. In this questionnaire, students were asked the question ‘is there any content from Year 10 that you feel is going to be hard for you to recall?’, and the subsequent responses

coded into the broad topic that they fall into according to the AQA specification. The results are shown below in Table 7.

Topic	Blank	Deserts	Coasts	Rivers	Hazards	Rainforests
Number	3	7	5	11	4	3

Table 7 Topics students identified that they would find hard to recall

The data from this question shows that the main areas students' highlighted concerns for were Rivers and Hot Deserts. This is broadly consistent with the incorrectly answered questions in the initial quiz, suggesting that students have a degree of self-awareness when it comes to knowledge of areas of weakness. This is interesting in attempts to ratify the existing evidence-base with the study carried out here. One of the suggested benefits of retrieval practice is that it allows students to identify areas of weakness on which to focus their revision efforts (Roediger and Karpicke, 2006; Karpicke, 2017). The students carried out the questionnaire after they had completed the multiple-choice quiz. Therefore, it is possible that completing the questions had allowed the students to identify areas that they found difficult, and they then subsequently reported these via the questionnaire, which would account for the broad agreement between the two sets of data as discussed above. However, as I do not have a control group to compare this data to, it is impossible to draw conclusions about this factor. It is also possible that the students would have identified the same areas of weakness if they had completed the questionnaire before the quiz.

The analysis of the outcome of the initial multiple-choice quiz was used to inform the retrieval practice intervention at the start, with questions about deserts and rivers the

initial focus, for example. The impact of the intervention on recall will now be explored.

4.2 Impact of Retrieval on Student Recall

Research question one regarding the influence that retrieval practice had on student recall is measured through analysis of the changes in student performance in the multiple-choice quizzes carried out in October, December, February, and May. At the start of data analysis, two students were excluded from analysis. This is because they had missed the initial test at the start of the study, and one or more of the subsequent tests so it was not possible to track changes over time. This leaves 24 students' data remaining to discuss.

4.2.1 Performance from Start to End of Intervention

Firstly, I have investigated how individual student scores changed from their initial test in October 2021, to the final one in May. The results of this are shown in

Figure 5.

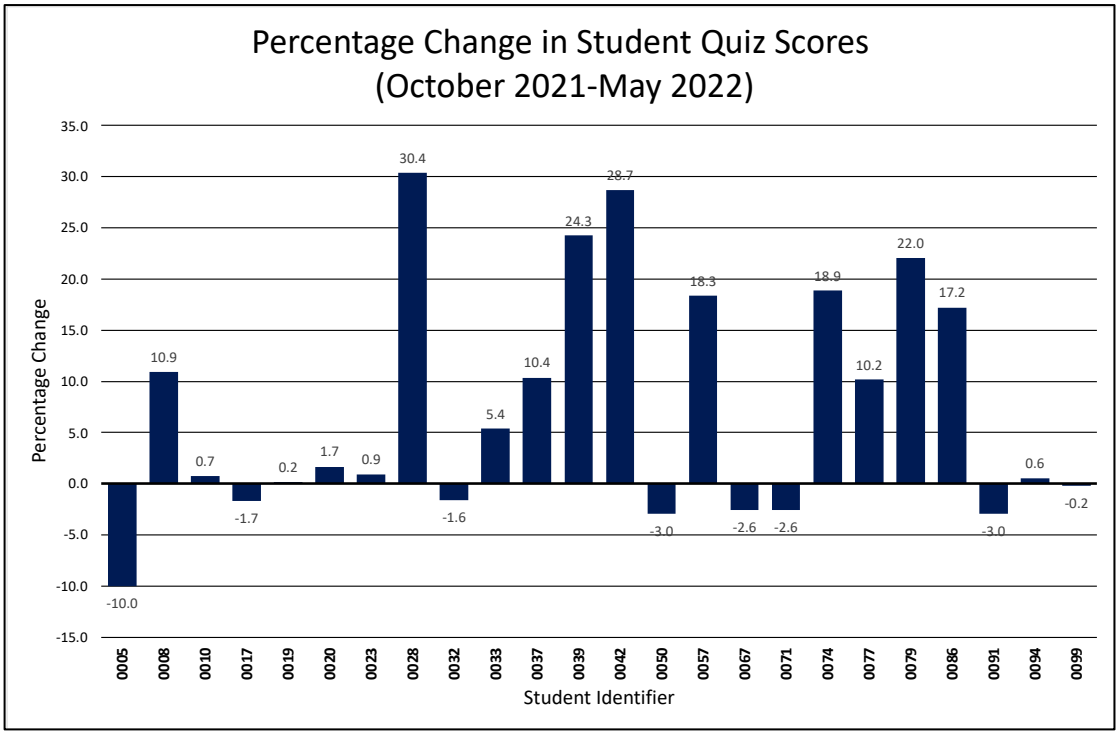

Figure 5 Graph showing percentage change in student scores

Figure 5 shows that 16 of the 24 students in the class (66%) have shown an increase in recall between the start and end of the intervention. This does mean that a further 33% of the students have decreased in scores over the test period. This warrants some exploration of the potential reasons, as it may indicate issues with the efficacy of the intervention.

Firstly, it should be noted that, except for student 0005, the decreases are less than 3%, equivalent to one or two incorrectly answered questions on the test, depending on the number of questions included. Furthermore, focusing on percentage changes obscures data on raw performance. For example, student 0032 decreased by 1.6% across the study, but their initial score was very high at 80%. This is also true for

students who showed very marginal percentage increase (e.g., student 0010 increased by 0.2% from a starting point of 75.9%). Although retrieval practice is widely credited with the ability to increase retention over time, it may be that this is less evident when students have very high initial recall. This is difficult to cross-reference with existing literature as most studies have not used core course materials that students will already have established memory of, but instead focus on arbitrary data such as word pairs (Carpenter, Pashler and Vul, 2006). However, it may also be that whilst engagement in retrieval practice hasn't *increased* the recall of students with high initial scores, it could have led to stronger long-term retention, or interruption of forgetting of knowledge that may have occurred without the intervention. This is also difficult to cross-reference with the literature, although given that most studies suggest that retrieval contributes positively to retention of information in the long-term memory (Gao *et al.*, 2016; Agarwal, Nunes and Blunt, 2021), this does at least seem likely.

A second factor to consider is that students' engagement with daily retrieval and performance on quizzes will have been highly variable across the study period. It therefore may be useful to look at individual case studies for students where performance has been either very poor, reflected in a large percentage decrease (Student 0005), or very strong, reflected in a large percentage increase relative to the study group (0028). A more detailed understanding of performance by these students can by elicited through qualitative data collected via the observations that were made of students by both myself and the LSA that collaborated with me on the project.

Next, I will review the available qualitative data for two students identified here with contrasting performance.

4.2.1.1 Student 0005

Student 0005 is both Pupil Premium and SEND and has a target grade of 5. Table 8 shows the performance of this student at each assessment point.

Student	October Score	December Score	February Score	May Score
0005	58.3%	50.0%	56.7%	48.3%

Table 8 Student 0005 data

This data shows that Student 0005's highest score was on the first multiple choice quiz, but that they did not decrease in a linear fashion. Overall, between the start and the end of the intervention, this student decreased by 10%, the greatest decrease in the class. As a result of having taught this student over a period of two years, I possess knowledge about their attitude, ability and background that helps to elicit reasons for the pattern identified. This is a benefit of practitioner research when compared to studies carried out by external researchers. The knowledge possessed by teachers about their students can be useful in providing insight into possible reasons for their performance on quizzes and tests.

As their teacher, I suggest that Student 0005 lacked confidence in attempting questions. Figure 6 is a photograph of work from this students' exercise book. When completing the retrieval activity at the start of the lesson, students were told to write down the question if they did not know the answer, and then record the correct answer when

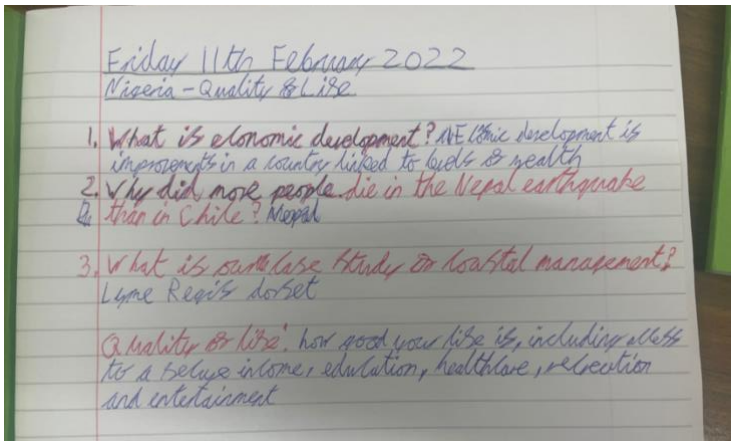

Figure 6 Extract of work for Student 0005

feedback was given to the class. Figure 6 shows that on this occasion, Student 0005 has written down each question and not attempted to answer them, instead recording the correct answer in different coloured pen when feedback was given. This is typical of the work completed by this student across the year. In terms of retrieval, the approach taken by this student is problematic, because it is the action of trying to recall information that strengthens memory traces (Roediger and Butler, 2011). It is likely that this student has not completed the questions because they found them difficult, and this is supported by observational information from the LSA collaborating with this study. The LSA reported that Student 0005 often did not respond to prompting to think of answers to the questions, claiming that the questions were 'too hard' and they 'couldn't remember'. The review of literature highlights that the difficulty of questions is key to successful retrieval, with more effortful retrieval associated with greater benefit to long-term retention (Bjork and Bjork, 2011; Karpicke, 2017).

Student 0005's reaction to the difficulty of the task highlights a problem with retrieval practice in the classroom, namely that its efficacy relies on students' being receptive to trying. However, difficulty of task is also important to think about from a pedagogical perspective – it is crucial that teachers optimise task difficulty whilst also ensuring that students can complete the work. The discussion here serves to indicate that successful and effective retrieval practice in the classroom is a multi-faceted problem that requires teachers to design and ask appropriately difficult questions to enable the retrieval effort influence to take place, but also requires students to commit to struggle in order to allow 'desirable difficulties' to take place.

4.2.1.2 Student 0028

Student 0028 is also Pupil Premium and has an EHCP (Educational Health Care Plan) as a result of their SEND needs. This student's target is a grade 2. Table 9 shows this student's score at each assessment point, and highlights that they have improved over time.

Student	October Score	December Score	February Score	May Score
0028	29.6%	70.0%	53.3%	60.0%

Table 9 Student 0028 data

This student works closely with the LSA, allowing me to elicit observational data that may account for the improvement shown by this student. Firstly, both myself and the LSA identified that this student did not complete all of the questions in the initial quiz – they only completed up to question 52, or half of the questions. This is likely a result of the large volume of reading that was required to answer the questions, problematic when the student has a low reading age and struggles with comprehension. The subsequent quizzes were much shorter, which meant that this student was able to attempt to answer all of the questions, which automatically increased their chance of scoring higher. I suggest this is at least part of the reason why this student has increased by such a large margin (40.4%) between October and December. I also asked the LSA to ask the student whether they would like the questions reading to them, which is an access arrangement that this student would have in the final exam as well.

Secondly, the LSA observed and reported that as the intervention progressed, student 0028 became more confident in attempting to answer the retrieval practice questions.

I do not have written evidence of this to present here, as this student verbally answered the questions with the LSA so that additional time was not wasted in writing. This student is also entitled to a scribe, so this further supports them in accessing the intervention to a similar level as their peers. Therefore, this shows that the student has attempted to engage with the intervention, and this may account for the improvement in their scores over time.

4.2.2 Changes Over Time

Students completed four multiple-choice quizzes over the period of the intervention, so performance can be assessed over time. However, only 17 of the students had sufficient attendance to complete all four of the data collection points. The changes in their performance over time will be explored here.

Figure 7 shows the percentage score achieved by each of the 17 individuals who completed the tests in October, December, February and May. Close inspection of this highlights that there are very few cases where students have incrementally improved (or regressed) with each subsequent test – in fact only student 0077 improved with each test.

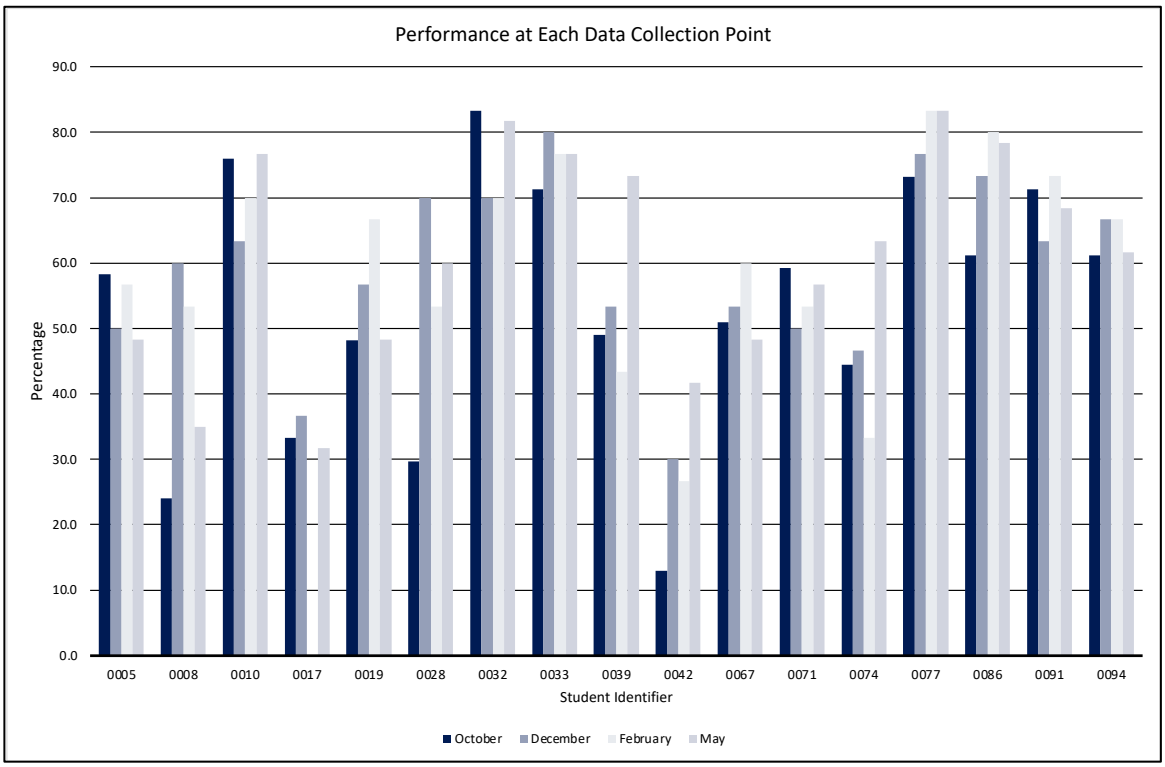

Figure 7 Graph showing performance of students at each data point

The data shown in Figure 7 suggests that the influence of retrieval practice on recall is not linear. However, on average, the majority of students increase in recall from the start to the end of the study. The class as a whole improved from an average score of 51.2% in October, to 59.9% in May. Given that the real-world application of the retrieval practice intervention implemented here would be to improve students' ability to recall knowledge needed in final GCSE examinations which are sat in May and June, it is perhaps not important exactly the rate at which students improve in recall, but that they do improve at all.

It is difficult to compare the data about changes in recall overtime with existing studies, as the literature reviewed has not identified research that has followed a class of students over a long period of time. Therefore, I do not have a benchmark with which to judge whether I may have expected recall to improve with each assessment

point. Similarly, it cannot be determined whether the increase in recall from 66% of the class is more or less than has been achieved in other studies, or whether a greater increase should have been expected or possible.

Another factor to consider in accounting for the non-linear change may be the vast range of individual differences that each student brings to the classroom environment (Agarwal *et al.,* 2017; Agarwal, Nunes and Blunt, 2021). These include those which can be systematically identified (such as SEND, EAL, target grades, gender, etc.), but also those which cannot, including engagement of the students on a day-to-day basis, the differing individuals that were absent on any one day, events that had happened during the school day prior to arrival in my classroom and any number of circumstances personal to each student. In addition to these variables, there are also extraneous variables particular to the classroom environment. This is important to note in terms of comparing this study to the data elicited from laboratory-based studies, as these variables will not exist in these circumstances (Roediger and Karpicke, 2006). It also raises issues in terms of generalising the data to other scenarios, as the variables experienced in my classroom will be different elsewhere, even in the same school.

4.2.3 Target Groups

Within the framework within which the school analyses student data, there are a number of 'target groups'. Given the suggested importance of individual differences on both student engagement and motivation in retrieval practice activities and their later performance in testing, I will now explore the data for a number of sub-groups

within the class. This will include disadvantaged students, SEND students, EAL students, those with target grades below a Grade 4, and those with target grades 5 and above.

4.2.3.1 Disadvantaged Students

The school identifies disadvantaged students as those who bring additional Pupil Premium funding to the school, whether because they come from a family with limited income, they are children of services families or they are looked after children. In my class, the disadvantaged students are identified as such because they are in receipt of free school meals. There are nine disadvantaged students in the class.

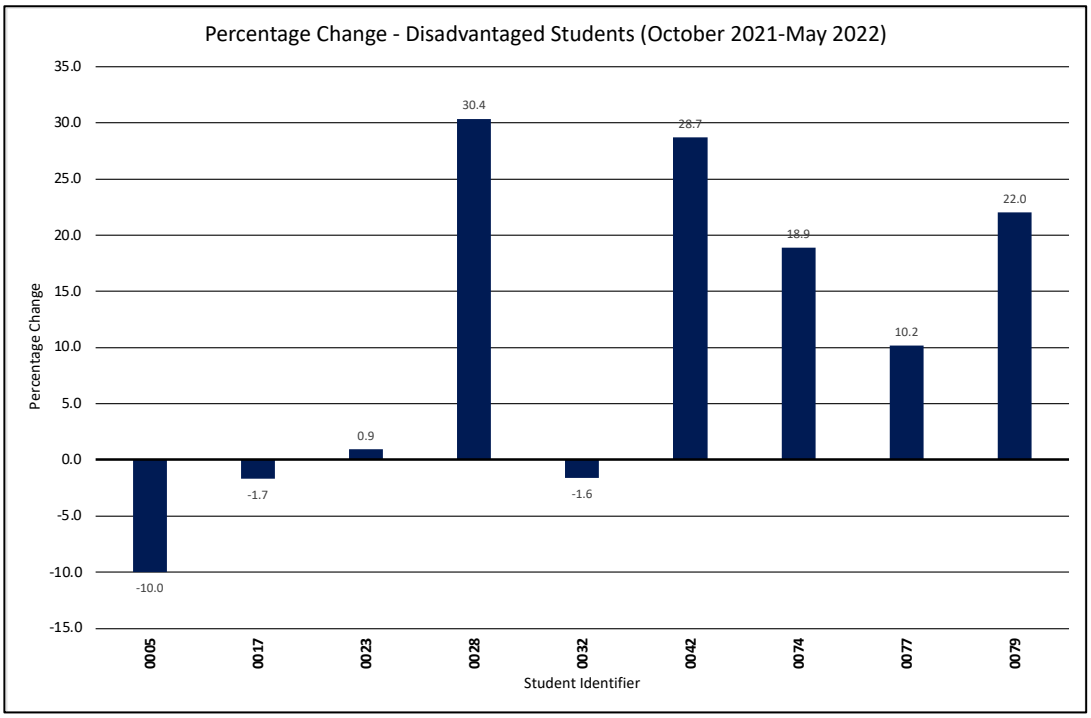

Figure 8 Graph showing percentage change for disadvantaged students

Figure 8 shows that six of the nine disadvantaged students have increased in recall over the intervention. At 66%, this is exactly in line with the average performance of the class as a whole. This is interesting given pervasive education rhetoric that states a

need for teachers to undertake activities to 'close the gap' between disadvantaged students and their non-disadvantaged peers.

Another factor to consider is the interplay that disadvantage has with other categorisations. For example, socioeconomic disadvantage is often coincident with SEND and EAL. This is something that will need consideration by each teacher if this intervention is to be successfully implemented on a wider scale.

4.2.3.2 Students with SEND

Students with SEND include those classified as either SEN Support (K) or with EHCPs (E) on the schools' SEND register. There are seven students with SEND in the class. Figure 9 is an extract from the whole-class data set showing the overall change in scores for the SEND students.

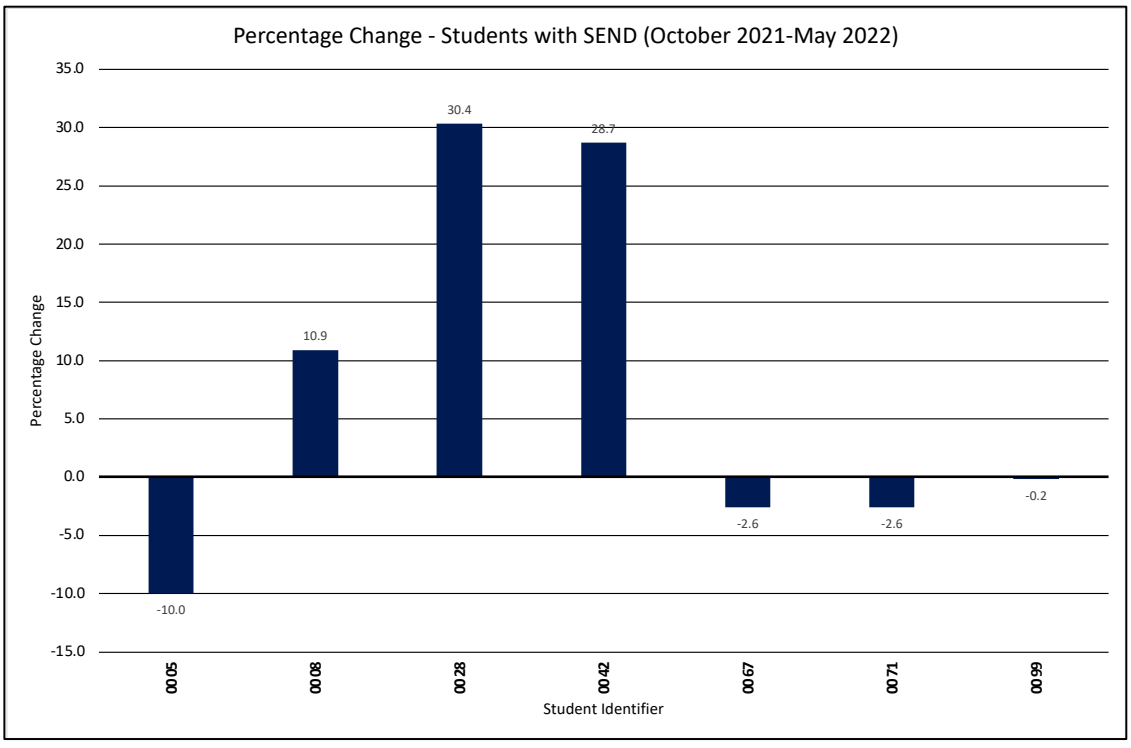

Figure 9 Graph showing percentage change for students with SEND

Analysis of the data for these students highlights that the student with the largest increase in scores and the student with the largest decrease are both in this group. The potential reasons for the performance of these students have been explored above. Whilst it is interesting to note that performance at each end of the spectrum falls within this group, I suggest that this does not necessarily indicate much about retrieval practice or the intervention itself. Firstly, the variation of needs that are encompassed under the umbrella of SEND in schools is very wide, and can include physical disabilities as well as issues with learning, memory and cognition. Student 0005 and 0028 will have very different needs catered for within their SEND diagnoses and therefore there is little utility in comparing their data. Furthermore, this study has not sought to explore the specific applications of retrieval practice for students with SEND, or the reasons for differences in their performance on quizzes, tests or final exams so it is not possible to provide an evidence-based commentary on this.

4.2.3.3 Students with EAL

Students with EAL are identified through the school's internal EAL register. There are five students with EAL in the class, and each has an EAL passport which details their proficiency in English according to the Department for Education's (DfE) EAL Proficiency Stages. The five EAL students are classified as either competent (D) or fluent (E) according to these stages. This means that they either operate across the curriculum without EAL support to the same level as first language speakers, or that they only require occasional support (DfE, 2020).

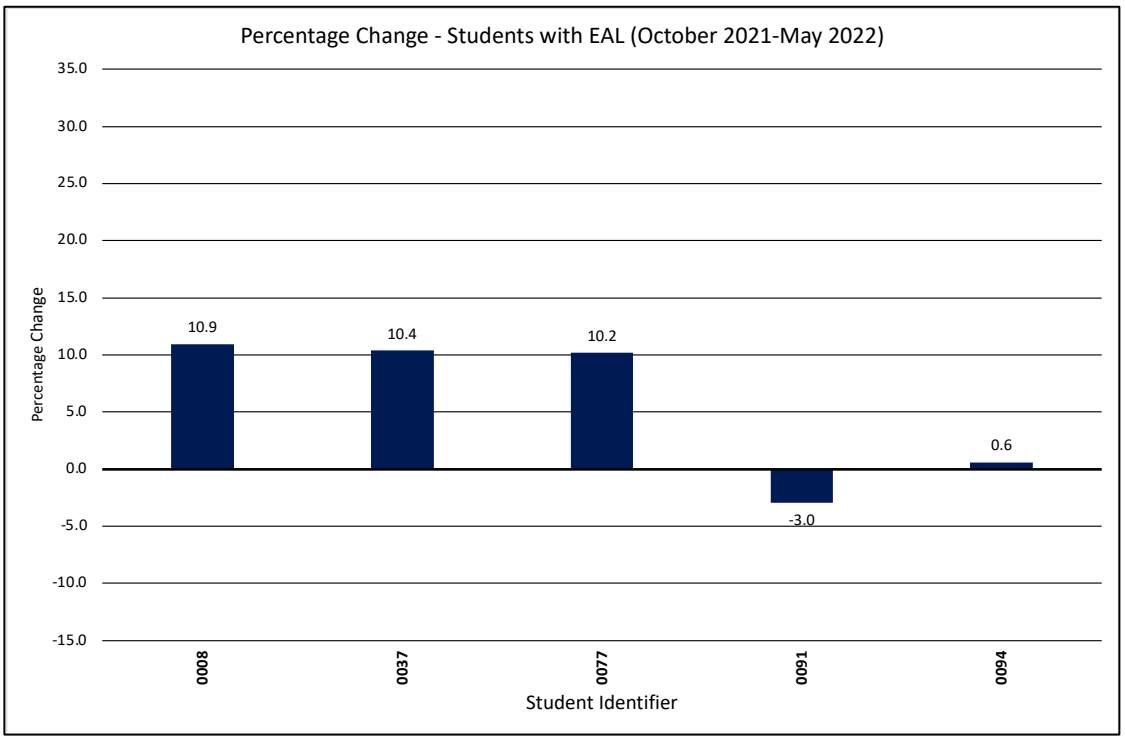

Figure 10 Graph showing percentage change for students with EAL

Figure 10 highlights that four out of the five EAL students increased in score from October to May, and the student that decreased was by a small margin of -3%. Closer inspection of the raw data shows that student 0091 who decreased over time had a high initial score of 71.3%, so a small decrease does not necessarily signal that retrieval practice has not been useful for this student.

It is also important to note that although English proficiency may be a minor influence in the case of these students, this may not be the case for students who are newer to English. The design of a retrieval practice intervention to support this type of student is an important consideration going forward. This is because for students who are less proficient in English, their ability to encode and recall information in an English-language medium is likely to be restricted. Therefore, teachers should be aware of this

and use any appropriate technique to overcome this. This will not be further explored in this study, as it is outside of the focus of this investigation.

4.2.3.4 Target Grade 4 and below

A 'standard pass' at GCSE is achieved with a grade 4. As a result, much data analysis that takes place within the school is focused on interventions to enable students to reach a grade 4. I have chosen to focus on students with a target grade 4 or below as a specific sub-group. This is because these are usually the less able students in the class, and as a result are more likely to struggle to access the exam questions to a level sufficient to achieve high marks. It was hoped that taking part in a retrieval practice intervention would improve the ability of students to recall key knowledge and also recognise key words in questions increasing their ability to pick up marks in the exam, making it more likely they would achieve their target grade. Figure 12 shows the percentage change in scores for these students.

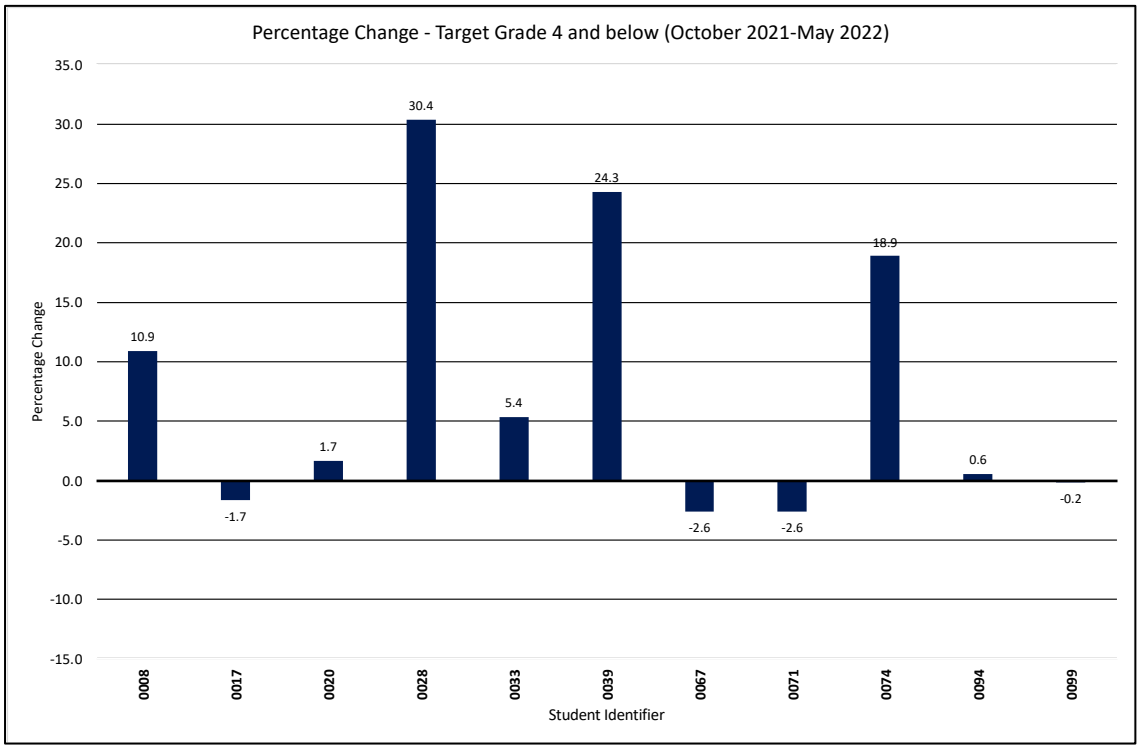

Figure 11 Graph showing percentage change for students with target grade 4 and below

Figure 11 shows that of the 11 students in this group, seven of them showed an improvement in scores overtime, and the remaining four showed only small decreases, equivalent to one or two incorrectly answered questions on the May quiz. This highlights that the intervention may have had some success in increasing recall for these students. The intention would be that this recall would translate to increased ability to apply knowledge to short answer questions in the final exam, as these are written in a similar way to the retrieval practice questions asked at the beginning of lessons.

4.2.3.5 Target Grade 5 and above

A 'strong pass' at GCSE is achieved with a grade 5. Students with target grades 5 and above are usually the more able students, and can include those with target grades at the top of the grading system. The highest target grade in the class taking part in this study is a grade 7. As a teacher, I know that for students to achieve grades 5 and above, they need to be answering the majority of short answer questions in the exam correctly, and also starting to pick up a larger majority of marks for higher tariff (6 and 9 mark) questions, in comparison to students who will achieve below a grade 5. Therefore, the aim of the retrieval practice intervention for these students was to ensure that they could recall appropriate detail to use in these questions. Figure 13 shows the data for this group.

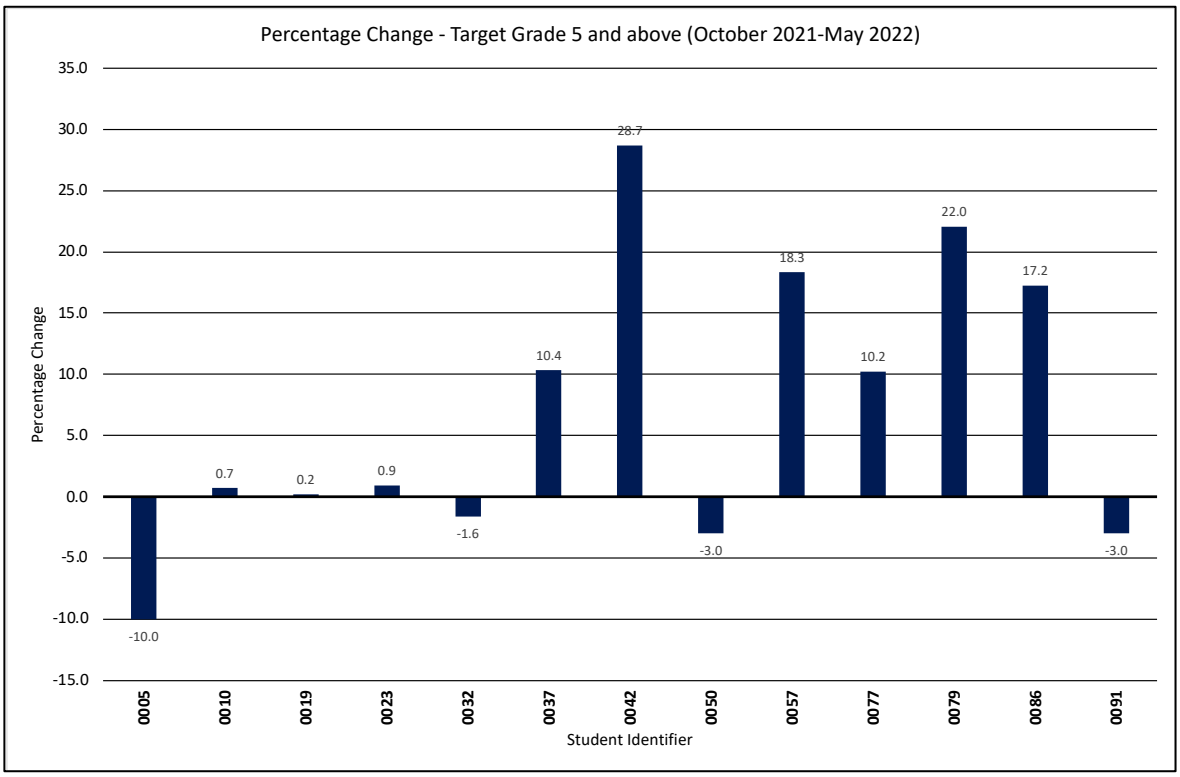

Figure 13 Graph showing percentage change for students with target grade 5 and above

Figure 13 shows that thirteen students have a target grade of 5 or above, and four of these show a decrease in scores over time. Students with the most marginal of increases and decreases also fall into this category. Again, this highlights that there is not necessarily a clear pattern in the influence retrieval practice has had on these students, although the majority of them have improved in recall overtime.

4.2.4 Explanations from Observation

The above exploration of the data collected in this study highlights that there is not necessarily a clear pattern in the influence that retrieval practice has had on student recall. Furthermore, there are a number of complicating factors that means it is not possible to separate the impact that the intervention may or may not have had on

students; including but not limited to any independent revision that students have done, homework completion and engagement with the intervention itself.

Over the period of the study, I observed behaviour in students that may have affected their results. For example, during one of the quizzes, one student was very clearly not reading the questions and trying to select the correct answer, but just ticking random answers. This was confirmed when I later spoke to this student, and they agreed that they had not tried to complete the quiz properly. This may have happened on more than one occasion, and therefore means that the data elicited by the quizzes is not reliable, nor a true reflection of the knowledge that students would be able to recall if completed properly.

A second observation made by both myself and the LSA in this study is that students did not always attempt to answer questions. I will not explore the reasons behind this, as it is beyond the scope of this study, but I posit that it may be due to a lack of confidence and unwillingness to record incorrect answers. This is problematic as it means that the students are not engaging fully with retrieval practice, and therefore the cognitive mechanisms by which it is suggested to work will not be able to take place. This includes, for example, the strengthening of memory traces in the elaborative retrieval theory (Carpenter, 2009) and the development of contextual cues that is key to the episodic context account (Karpicke *et al.,* 2014) of retrieval practice. I suggest that this could be deemed a failure of the intervention itself, as students have not fully engaged with it and therefore it cannot be as successful as if all students completed the questions at all times.

These observations further serve to highlight that the success of this particular intervention is in part governed by the engagement and willingness of the students in question to attempt the tasks. This is a variable that is difficult, if not impossible, to control in the classroom. I suggest that a way to start to overcome this is to fully explain to students why they are being asked to carry out retrieval practice in each lesson, therefore increasing their metacognitive awareness of the benefits of the technique to their own learning (Karpicke and Grimaldi, 2012), and encouraging them to engage with the technique.

4.2.6 Review of Intervention for Recall

Overall, there has been an improvement in scores on the quizzes for two-thirds of the class, but it is not possible to determine the extent to which the retrieval practice intervention has been responsible for this, compared to other factors that could increase or decrease recall.

Despite these criticisms, the intervention was very easy to implement in a classroom context. The questions asked at the start of the lessons were easy to plan and the majority of the students in the class willingly completed them, and then contributed answers to class discussion. Furthermore, the questions are now planned and can be used again with limited revision. Therefore, I suggest that this is an easy way to continue to encourage students to take part in revision activities in a classroom environment, regardless of whether the intervention itself has a large measurable impact on student recall.

Furthermore, the literature on retrieval practice from both laboratory and classroom-based studies indicates large positive effects of retrieval practice in comparison to other study methods, suggesting that whilst the data in this study cannot be used to draw substantive conclusions by itself, it does provide further support to the already extensive evidence-base.

4.2.7 Important Aspects of the Intervention

The current evidence-base highlights important factors to consider when implementing retrieval practice in the classroom. Many of these were utilised in this study, and I suggest should be maintained in the wider use of the intervention in the school.

Firstly, the provision of feedback immediately following the attempts by students to answer the questions is important (Karpicke, 2017). Feedback is necessary in overcoming issues with individual performance, for example where students have got answer wrong or not answered it at all. Feedback was given in this study by asking students to contribute answers, and then displaying the correct answer on the board. Students were then encouraged to write down the correct answer in a different colour pen in their exercise book. This allowed me to identify mistakes made and where lures had been assimilated into long-term memory (Butler and Roediger, 2008; Marsh and Cantor, 2014; Butler *et al.,* 2017), and also ensure that students processed the answers by recording them (Marsh and Cantor, 2014).

A further consideration that was achieved in this study was the need to ensure that retrieval practice remained low-stakes to mitigate the negative impacts on memory caused by anxiety in relation to testing (Khanna, 2015). This was achieved as I did not record the scores of students on the questions at the start of lessons, and the scores for the quizzes have only been used in this study, not for any grading in school.

4.3 Impact of Retrieval on Student Confidence

Changes in student confidence are assessed through comparison of the pre- and post-intervention questionnaires. Data from limited interviews and observations also provides some insight.

4.3.1 Confidence from Start to End of Intervention

In both the pre- and post-intervention questionnaires students were asked how confident they felt in applying knowledge learned in the previous year to answering exam questions. This question was asked rather than just focusing on confidence in recall generally because students would need to apply learning to an exam at the end of the course. Figure 14 shows the comparison of the scores given by students on a Likert scale.

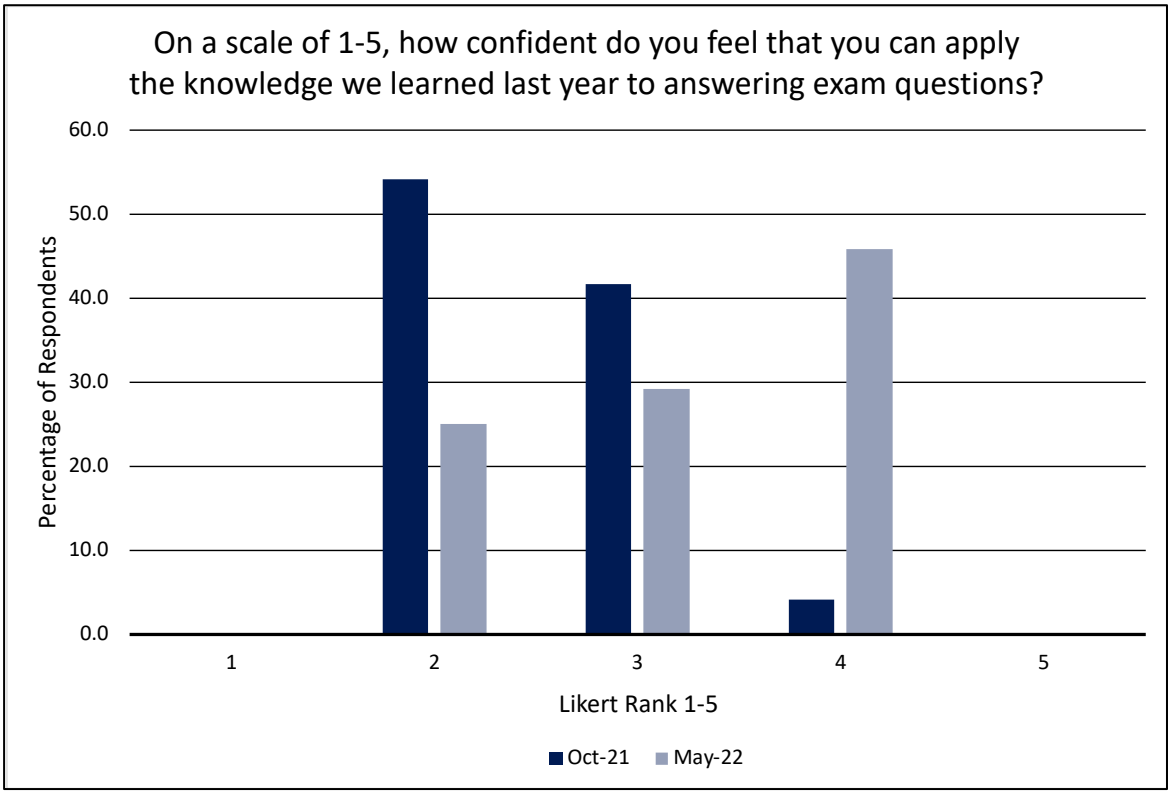

Figure 14 Graph showing changes in reported student confidence over time

Figure 14 demonstrates that in October, 54.2% of students ranked their confidence in applying knowledge to exam questions as 2 out of 5. The highest ranking was a 4, but only one student selected this score, highlighting that the confidence of most students was low. Although the main aim of the intervention in this study was to increase the amount of knowledge that students could recall in relation to questions in their GCSE exam, increasing student confidence was also important, particularly in relation to reducing test anxiety both for the in-class quizzes and the final exams. The data above shows that the intervention has demonstrated success in improving confidence, as in May, the majority of students (45.8%) reported confidence as 4 out of 5, an increase from the beginning of the intervention. However, it is also important to note that there were still 25% of students reporting their confidence as 2 out of 5. Therefore, there were still a significant number of students who did not feel their confidence had improved, highlighting that the intervention has not been a complete success.

However, this data relies on students reporting their own confidence accurately, and this cannot be guaranteed. A further problem is that the questionnaires were completed anonymously by students. This means that I cannot compare the changes in confidence for the same target groups as I have explored changes in recall (SEND, EAL, disadvantaged and different target grades). In addition, there is little available literature on confidence with which to compare my findings to. The main area of comparison pertains to test anxiety and suggests that taking part in low-stakes retrieval practice on a regular basis reduces the anxiety felt by students (Leeming, 2002; Agarwal *et al.,* 2014; Karpicke, 2017). It is possible that reduced anxiety leads to students feeling more confident, but I do not have the data to support this claim here.

4.3.3 Interview Data

The two short interviews conducted in May 2022 provide some insight into student confidence. The first student interviewed was positive about the retrieval practice intervention, stating that 'it made us think about all the different topics all the time'. This student also highlighted the areas of the course they had found retrieval practice most useful for – 'it has been really useful for case study facts and for key terms like hydraulic action'. This student said, 'I definitely feel more confident that I can remember things for the exam now', which highlights that the intervention has had some success in improving confidence for this student.

The second student interviewed offered similar points of view. They stated that 'we actually do a similar thing in Maths and Science as well and its really helpful' and 'I wish they would do it in English as well because it's helped me so much in

remembering the topics'. Again, this statement shows that this student has found retrieval practice useful and suggests that carrying out similar interventions around the school increases student buy-in.

Despite the positive comments outlined here, it is important to remember that this data has come from only a small sub-section of students in the class. Furthermore, these students volunteered to talk to me about retrieval practice, and therefore may have been more likely to have positive views. In addition, there is always power relationships at play in interviews, but particularly in the case of a teacher interviewing their students (Eder and Fingerson, 2003). This means that the data elicited here is not only limited in scope, but may also be unreliable, as students may have just said what they thought I would want to hear. Therefore, this is an area of this study that could be extended further, perhaps by selecting the students to interview more systematically (to include, for example, students with SEND and differing target grades), and having an external person complete the interviews so as to overcome some of the issues with power dynamics.

4.3.4 Review of Intervention for Confidence

Overall, the available data suggests that the intervention carried out here has had some success in improving student confidence in relation to recall of material linked to answering exam questions. However, the data is limited to individual questions on two questionnaires and there is limited research evidence to compare this to. I have also not been able to account for any additional factors that may have increased confidence overtime, such as practice of exam questions in lessons, independent

revision or homework completion. This means that it is not possible to draw substantive conclusions as to the influence that the retrieval practice intervention has had on student confidence. However, given that it is a simple intervention to implement in the classroom, I suggest that it is worthwhile continuing with it, in order to benefit from any potential influence on student confidence.

4.3 Summary

Overall, the data elicited in this study is not sufficient for me to draw conclusions that can be generalised outside of this study. However, my findings demonstrate that the intervention in this study has had some utility in increasing both the recall and confidence of a majority of the students in the study group.

5. Conclusion

5.1 Conclusions to Research Questions

5.1.1 In what ways does retrieval practice influence Year 11 students' recall?

The first research question in this study was regarding the influence that the retrieval practice intervention had on students' recall. The data explored above shows that 66% of the class increased in scores in multiple-choice quizzes from the start to the end of the intervention. Of the remaining 33% of students who decreased in scores, most of these were only by a small margin. This suggests that the influence of the intervention has been to increase the recall of knowledge exhibited by these students. This finding is supported by the exploration of evidence in the literature review which suggests retrieval practice has a large positive impact on student recall.

5.1.2 Does taking part in retrieval practice influence students' confidence in recall?

The second research question intended to determine whether taking part in retrieval practice has an influence on students' confidence. The data explored here suggests that taking part in the intervention has led to an increase in student confidence in terms of applying knowledge to exam questions. This is important as answering exam questions accurately is the way in which students gain marks to achieve a grade in the GCSE course as a whole, and therefore increase confidence can be seen as a success of the intervention. This aspect of the study adds a different dimension than that explored in the existing literature, which has focused on test anxiety rather than confidence in recall.

5.1.3 Limitations of the Data and Study

As explored in the previous chapter, the data produced in this study is limited in its extent, having been focused on one class and not supported with data from control groups. Therefore, it is impossible to determine if any of the influences on confidence and recall explored have occurred due to the intervention carried out, or any number of variables that could not be controlled as part of this study. Furthermore, the study completed here was designed specifically to strengthen the recall of knowledge from Paper 1 of GCSE Geography. This means that the intervention itself as well as the findings may not be transferrable to other contexts.

5.2 Implications

Despite the limitations that have been explained, this intervention still has some utility for use in the school going forward. Firstly, the intervention has been designed taking into account a wide range of research. This means that a number of issues associated with retrieval practice (for example, test anxiety, rote learning, need for feedback) explored in the literature review have already been considered, and solutions identified. Therefore, the intervention itself has been informed by the available evidence and can be easily implemented. Secondly, my experiences this year show that the intervention can easily form part of day-to-day classroom teaching, without being onerous to plan or taking too much time out of lessons. This is important as teachers have limited time in which to teach core course content and also revise material. An effective retrieval practice intervention such as that detailed here helps save time by embedding revision in every lesson. Finally, if the intervention is used more widely than just in my classroom, it is hoped that students will further realise

the benefit of engaging with the technique, increasing their metacognitive awareness of their areas of weakness and encouraging them to engage with the 'desirable difficulty' of effortful retrieval.

5.3 Further Research

There are a number of areas that could benefit from further research to improve the intervention in both my school and on a wider scale. This includes attempting to find a way to encourage students to embrace the challenge of answering questions that are difficult so that they benefit from effortful retrieval and the associated increase in strength of memory traces. A second area for further research and development of the intervention is to find a way to increase student motivation to complete retrieval practice in their own revision outside of the classroom, given the extensive evidence which suggests this method is superior to other widely used strategies such as re-reading and re-study for increasing recall.

References

Abbott, E.E. (1909) On the Analysis of the Factor of Recall in Learning. https://www.teachertoolkit.co.uk/wp-content/uploads/2020/09/abott1909.pdf (Accessed: 14th May 2022)

Agarwal, P. (2019). 'Retrieval Practice & Bloom's Taxonomy: Do Students Need Fact Knowledge Before Higher Order Learning?' *Journal of Educational Psychology*, 111, 189-209. https://doi.org/10.1037/edu0000282.

Agarwal, P.K., Nunes, L.D. and Blunt, J.R (2021) 'Retrieval Practice Consistently Benefits Student Learning: a Systematic Review of Applied Research in Schools and Classrooms.' *Educational Psychology Review,* 33:1409-1453 doi: https://doi.org/10.1007/s10648-021-09595-9 1

Agarwal, PK., *et al.* (2014) 'Classroom-based programs of retrieval practice reduce middle school and high school students' test anxiety.' *Journal of Applied Research in Memory and Cognition,* 3, pp. 131-139 doi: http://dx.doi.org/10.1016/j.jarmac.2014.07.002

Agarwal, PK., *et al*. (2020) *How to use retrieval practice to improve learning*. Available online: retrievalpractice.org/library (Accessed: 10th June 2022)

Alferink, L.A. and Farmer-Dougan, V. (2010) 'Brain-(not) Based Education: Dangers of Misunderstanding and Misapplication of Neuroscience Research.' *Exceptionality,* 18:1, 42-52 doi: https://doi.org/10.1080/09362830903462573

Atkinson, R.C. and Shiffrin, R.M. (1968) 'Human Memory: A Proposed System and its Processes'. *Psychology of Learning and Motivation,* 2, pp.89-195 doi: https://doi.org/10.1016/S0079-7421(08)60422-3

BERA (2018) *Ethical Guidelines for Educational Research 4th Edition*. Available at: https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-2018-online (Accessed: 14th August 2021).

Bjork, E. L., and Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M.

Hough, J. R. Pomerantz (Eds.) & FABBS Foundation, *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Bjork, R, (2012) Gocognitive: 'Robert Bjork – using our memory shapes our memory YouTube [Video] 13 July 2012. Available online: https://www.youtube.com/watch?v=69VPjsgm-E0 (Accessed: 31st May 2022)

Bjork, R.A. (1994) Memory and Metamemory Considerations in the Training of Human Beings. In J. Metcalfe (Eds) Metacognition: Knowing and Knowing (pp.185-205) Cambridge, MA: MIT Press.

Bjork, R.A., and Bjork, E.L. (1992) A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn and R. Shiffrin (Eds.) *From Learning Processes to Cognitive Processes: Essays in Memory of William R Estes* (Vol. 2, pp.35-67) Hillsdale, NJ: Erlbaum

Blainey, K. *et al* (2020) 2020 White Paper. *The Impact of Lockdown on Children's Education: A Nationwide Analysis*. RS Assessment From Hodder Education (November 2020) Available online: https://www.risingstars-uk.com/media/Rising-Stars/Assessment/Whitepapers/RS_Assessment_white_paper_1.pdf (Accessed: 20th June 2022)

Carpenter, S.K. (2009). 'Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval.' *Journal of Experimental Psychology: Learning, Memory & Cognition*, 35, 1563–1569 doi: https://doi.org/10.1037/a0017021

Carpenter, SK., Pashler, H. and Vul, E. (2006) 'What types of learning are enhanced by a cued recall test?' *Psychonomic Bulletin & Review,* 12 (5), pp. 826-830 doi: https://doi.org/10.3758/BF03194004

De Weerdte, F., Desoete, A., and Royers, H. (2013) 'Working Memory in Children with Reading Disabilities and/or Mathematical Disabilities'. *Journal of Learning Disabilities,* 46(5), pp. 461-472 doi: https://doi.org/10.1177/0022219412455238

Department for Education (DfE) (2019) *Early Career Framework* DFE-00015-2019. Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/978358/Early-Career_Framework_April_2021.pdf (Accessed: 10th September 2021)

Eder, D. and Fingerson, L. (2003) 'Interviewing children and adolescents', in JA, H. and JF, G. (eds) *Inside Interviewing: New Lenses, New Concerns*. California: Thousand Oaks, pp. 33–53.

Education Endowment Foundation (EEF) (2020) The Impact of COVID-19 on Learning; A review of the evidence. Evidence Review. Available online: https://educationendowmentfoundation.org.uk/guidance-for-teachers/covid-19-resources/best-evidence-on-impact-of-covid-19-on-pupil-attainment (Accessed: 12th June 2022)

Education Endowment Foundation (EEF) (2021) *Cognitive Science Approaches in the Classroom: A Review of the Evidence.* Birmingham, England. Available online: https://educationendowmentfoundation.org.uk/public/files/Publications/Cognitive_science_approaches_in_the_classroom_-_A_review_of_the_evidence.pdf (Accessed: 10th August 2022)

Endres, T. *et al.* (2020) 'It matters how to recall – task differences in retrieval practice.' *Instructional Science,* 48, pp. 699-728 doi: https://doi.org/0.1007/s11251-020-09526-1

Gao, C., Rosburg, T., Hou, M. *et al.* (2016) 'The role of retrieval mode and retrieval orientation in retrieval practice: insights from comparing recognition memory testing formats and restudying.' *Cognitive Affective & Behavioural Neuroscience* **16,** 977–990 doi: https://doi.org/10.3758/s13415-016-0446-z

Gates, A.I. (1917). 'Recitation as a factor in memorizing.' *Archives of Psychology, 40,* pp. 104.

Gurung, RAR., and Burns, K. (2018) 'Putting evidence-based claims to the test: A multi-site classroom study of retrieval practice and spaced practice.' *Applied Cognitive Psychology,* 33(5), pp. 732-743 doi: **https://doi.org/10.1002/acp.3507**

Jones, K. (2019) *Retrieval Practice: Research and Resources for Every Classroom*. John Catt Educational Ltd. UK

Jones, K. (2021a) *Retrieval Practice: Implementing, Embedding and Reflecting*. John Catt Educational Ltd. UK

Jones, K. (2021b) *Retrieval Practice: Resource Guide: Ideas & Activities for the Classroom*. John Catt Educational Ltd. UK

Kang, SHK., and Pashler, H. (2014) 'Is the benefit of retrieval practice modulated by motivation?' *Journal of Applied Research in Memory and Cognition*, 3, pp.183-188 doi: https://doi.org/10.1016/j.jarmac.2014.05.006

Karpicke, J. D. (2017). Retrieval-based learning: A decade of progress. In J. T. Wixted (Ed.), Cognitive psychology of memory, Vol. 2 of learning and memory: A comprehensive reference (J. H. Byrne, Series Ed.)(pp. 487–514). Oxford: Academic Press

Karpicke, J. D., & Zaromb, F. M. (2010). 'Retrieval mode distinguishes the testing effect from the generation effect.' *Journal of Memory and Language, 62*(3), 227–239. https://doi.org/10.1016/j.jml.2009.11.010

Karpicke, J.D and Grimaldi, P.J. (2012) 'Retrieval-Based Learning: A Perspective for Enhancing Meaningful Learning.' *Educational Psychology Review,* 24:401-418 doi: https://doi.org/10.1007/s10648-012-9202-2

Karpicke, J.D. (2012) 'Retrieval-Based Learning: Active Retrieval Promotes Meaningful Learning.' *Current Directions in Psychological Science,* 21(3), pp.157-163 doi: https://doi.org/10.1177/0963721412443552

Karpicke, J.D., Butler, A.C, Roediger, H.L III (2009) 'Metacognitive Strategies in Student Learning: Do Students Practise Retrieval when they Study on their Own? *Memory,* 17(4), pp.471-479 doi: https://doi.org/10.1080/09658210802647009

Karpicke, J.D., Lehman, M., and Aue, W.R., (2014). Retrieval-based learning: an episodic context account. In: Ross, B.H. (Ed.), *Psychology of Learning and Motivation,* vol. 61. Elsevier Academic Press, San Diego, CA, pp. 237–284.

Khanna, MM. (2015) 'Ungraded Pop-Quizzes: Test-Enhanced Learning Without All the Anxiety.' *Teaching of Psychology*, 42(2), 174-178 doi: https://doi.org/10.1177/0098628315573144

Kolers, P. A., and Roediger, H. L. III. (1984). 'Procedures of mind.' *Journal of Verbal Learning & Verbal Behavior, 23*(4), 425-449. https://doi.org/10.1016/S0022-5371(84)90282-2

Leeming, F. C. (2002). 'The exam-a-day procedure improves performance in psychology classes'. *Teaching of Psychology, 29*(3), 210–212. https://doi.org/10.1207/S15328023TOP2903_06

Little, J.L., *et al.* (2012). 'Multiple-choice tests exonerated, at least of some charges: fostering test-induced learning and avoiding test-induced forgetting.' *Psychological Science*, 23 (11), 1337–1344. http://dx.doi.org/10.1177/0956797612443370.

Maldonado, JE., and De Witte, K. (2020) 'The effect of school closures on standardised student test outcomes.' *British Educational Research Journal*, 48 (1), pp. 49-94 doi: https://doi.org/10.1002/berj.3754

Marsh, E.J. and Cantor, A.D. (2014). 'Learning from the test: Dos and don'ts for using multiple-choice tests.' In M. McDaniel, R. Frey, S. Fitzpatrick, & H.L. Roediger (Eds), Integrating cognitive science with innovative teaching in STEM disciplines. doi: http://dx.doi.org/10.7936/K7Z60KZK

McDaniel, M.A. *et al.* (2007) 'Testing the Testing Effect in the Classroom'. *European Journal of Cognitive Psychology,* 19(4-5), pp.494-513 doi: https://doi.org/10.1080/09541440701326154

McDermott, K.B. *et al.* (2014) 'Both Multiple-Choice and Short-Answer Quizzes Enhance Later Exam Performance in Middle and High School Classes.' *Journal of Experimental Psychology: Applied,* 20(1), pp.3-21 doi: https://doi.org/10.1037/xap0000004

Morris, C.D, Bransford, J.D., and Franks, J.J (1977) 'Level of processing versus transfer appropriate processing'. *Journal of Verbal Learning and Verbal Behaviour,* 16(5), pp.519-533 doi: https://doi.org/10.1016/S0022-5371(77)80016-9

O'Donnell, C. L. (2008) 'Defining, Conceptualizing, and Measuring Fidelity of Implementation and Its Relationship to Outcomes in K–12 Curriculum Intervention Research', *Review of Educational Research*, 78(1), pp. 33–84. doi: https://doi.org/10.3102/0034654307313793

Ofqual (2018) *Guidance - Get the facts: GCSE Reform* (Updated 28 January 2018) Available online: https://www.gov.uk/government/publications/get-the-facts-gcse-and-a-level-reform/get-the-facts-gcse-reform (Accessed: 31st May 2022)

Ofsted (2019) *The Education Inspection Framework.* (England) Reference no: 190015. Available online: https://feweek.co.uk/wp-content/uploads/2019/05/Education_inspection_framework.pdf (Accessed: 10th September 2021)

Ofsted (2022) *Education Recovery in Schools: Spring 2022*. Oftsed Research and Analysis (4 April 2022). Available online: https://www.gov.uk/government/publications/education-recovery-in-schools-spring-2022/education-recovery-in-schools-spring-2022 (Accessed: 12th June 2022)

Pan, S.C and Agarwal, P.K (2020) *Retrieval Practice and Transfer of Learning: Fostering Students' Transfer of Knowledge*. Available online: http://pdf.retrievalpractice.org/TransferGuide.pdf (Accessed: 10th June 2022)

Pan, SC., and Agarwal, PK. (2020) *Retrieval practice and transfer of learning: Fostering students' application of knowledge.* Available online: retrievalpractice.org/library [Accessed 10th June 2022]

Perry, T. *et al.* (2021). Cognitive Science in the Classroom. London: Education Endowment Foundation (EEF). The report is available from: https://educationendowmentfoundation.org.uk/evidence-summaries/evidencereviews/cognitive-science-approaches-in-the-classroom/ (Accessed: 10th August 2021)

Roediger, H. L. (1990). 'Implicit memory: Retention without remembering.' *American Psychologist, 45*(9), pp.1043–1056. https://doi.org/10.1037/0003-066X.45.9.1043

Roediger, H. L. and Karpicke, J. D. (2006) 'Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention', *Psychological Science*, 17(3), pp. 249–255. doi: https://doi.org/10.1111/j.1467-9280.2006.01693.x.

Roediger, H.L. *et al*. (2011) 'Test-Enhanced Learning in the Classroom: Long-Term Improvements from Quizzing.' *Journal of Experimental Psychology: Applied*, Vol. 17, No. 4, 382-395 doi: https://doi.org/10.1037/a0026252

Roediger, HL., and Butler, AC. (2011). 'The critical role of retrieval practice in long-term retention.' *Trends in Cognitive Sciences* 15:1, pp.20-27 doi: https://doi.org/10.1016/j.tics.2010.09.003

Rowland, C. A. (2014). 'The effect of testing versus restudy on retention: A meta-analytic review of the testing effect.' *Psychological Bulletin, 140*(6), 1432-1463. https://doi.org/10.1037/a0037559

Rowley, T. and McCrudden, M.T (2020) 'Retrieval practice and retention of course content in a middle school science classroom.' *Applied Cognitive Psycholo*gy, 34:1510-1515 doi: https://doi.org/10.1002/acp.3710

Steele, D., 2011. Does test-taking help students learn. *New York Times A*, *24*.

Storey, N., and Zhang, Q. (2021) A Meta-Analysis of COVID Learning Loss [WWW document] doi: https://doi.org/10.35542/osf.io/qekw2

Szpunar, K. K., Khan, N. Y., and Schacter, D. L. (2013). 'Interpolated memory tests reduce mind wandering and improve learning of online lectures.' *PNAS Proceedings of the National Academy of Sciences of the United States of America, 110*(16), pp.6313–6317. https://doi.org/10.1073/pnas.1221764110

Watkins, O. C., and Watkins, M. J. (1975). 'Build-up of Proactive Inhibition as a Cue-Overload Effect.' *Journal of Experimental Psychology: Human Learning and Memory, 1*(4), 442–452. https://doi.org/10.1037/0278-7393.1.4.442

Appendices

Appendix 1: Multiple-Choice Quiz

The following is an extract from the first multiple-choice quiz. All following quizzes followed the same structure.

1. **"Where plates move towards each other, creating fold mountains". Which is the correct definition for this?**
   a. Constructive
   b. Destructive subduction
   c. Destructive collision
   d. Conservative

2. **Which of the following is a primary effect of the earthquake in Chile?**
   a. 12,000 people injured
   b. Coastal towns destroyed by tsunami waves
   c. 200 people injured
   d. Fires

3. **Which of the following is a secondary effect of the earthquake in Nepal?**
   a. 9000 people died
   b. Houses destroyed
   c. Avalanches on Mt Everest
   d. Tsunami

4. **Which types of plate meet at a destructive subduction boundary?**
   a. Oceanic – oceanic
   b. Continental – continental
   c. Oceanic – continental

5. **What is the name of the point on the Earth's surface, exactly above where the earthquake originates?**
   a. Focus
   b. Epicentre
   c. Shockwave
   d. Origin

6. **What temperature does water need to be for a tropical storm to form?**
   a. 24°C
   b. 28°C
   c. 27°F
   d. 27°C

7. **Which season do most tropical storms happen?**
   a. Summer and Autumn
   b. Spring and Summer
   c. Summer and Winter
   d. Winter and Spring

8. **What happens when a tropical storm reaches land?**
   a. It gets stronger
   b. It loses energy
   c. It picks up more wind
   d. It gets bigger

9. **Which of the following is an example of extreme weather in the UK?**
   a. A thunderstorm in Manchester
   b. Snow in Northern Scotland
   c. Drought in the North
   d. Rain in Farnborough

10. **Which of the following definitions best describes the following statement? "A natural cause of climate change that occurs when the Earth's orbit changes from being circular to elliptical over a period of 100,000 years.**
    a. Greenhouse effect
    b. Obliquity
    c. Sun spots
    d. Eccentricity

Appendix 2: Questionnaire

GCSE Geography Revision Questionnaire

1. On a scale where 1 is **not at al**l and 5 is **very**, how confident do you feel that you can remember the content we covered last year?

   1 2 3 4 5

2. On a scale where 1 is not at all and 5 is very, how confident do you feel that you can apply the knowledge we learned last year to answering exam questions

   1 2 3 4 5

3. Is there any content from Year 10 that you feel is going to be particularly hard for you remember? What is it?

   _____

4. On a scale where 1 is not at all and 5 is very, how confident do you feel that you know how revise?

   1 2 3 4 5

5. Have you heard of retrieval practice?

   Yes No

6. Do you know how to use retrieval practice to help you remember material you have learned in the past?

   Yes No

7. Below, please write down any methods of revision that you intend to use to help you prepare for your GCSE Geography exam.

8. Is there anything that you feel you will need particular help with to prepare you for your GCSE exam?

Appendix 3: Example of Question-Level Records

	A: TECTONIC HAZARDS					A: WEATHER HAZARDS					B: TRF					B: HOT DESERTS					C: COASTS					C: RIVERS						
Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Total	%
0005	N	Y	Y	N	Y	N	N	Y	N	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	N	N	N	N	N	N	N	Y	N	15	50.0
0008	N	Y	N	Y	Y	Y	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	N	Y	N	N	N	N	Y	18	60.0
0010	Y	Y	N	Y	Y	Y	N	Y	Y	N	Y	Y	N	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	N	N	N	Y	19	63.3
0017	N	N	N	N	Y	N	N	Y	Y	N	N	Y	Y	N	N	N	N	Y	Y	N	N	Y	N	N	Y	N	Y	N	N	Y	11	36.7
0019	N	Y	N	N	Y	Y	N	Y	N	N	Y	Y	N	Y	N	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N	Y	Y	N	N	17	56.7
0020																															0	0.0
0023	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	N	N	N	N	18	60.0
0028	Y	Y	N	N	N	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	N	Y	21	70.0
0032	N	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	N	Y	N	N	Y	Y	Y	N	Y	21	70.0
0033	Y	Y	Y	Y	Y	Y	N	Y	Y	N	N	Y	N	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	24	80.0
0037	Y	N	Y	Y	Y	N	N	Y	N	N	N	Y	N	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	20	66.7
0039	N	Y	Y	Y	Y	Y	N	N	N	Y	Y	N	N	Y	N	Y	Y	N	Y	N	Y	N	Y	N	N	Y	N	Y	N	Y	16	53.3
0042	N	Y	N	N	Y	N	N	N	N	N	Y	Y	N	Y	N	N	N	N	Y	N	N	N	N	N	Y	N	N	N	Y	Y	9	30.0
0046	N	Y	N	Y	N	Y	Y	N	Y	Y	Y	Y	N	Y	N	N	Y	N	Y	Y	N	N	N	Y	Y	N	N	N	N	N	14	46.7
0050																															0	0.0
0057	N	Y	Y	Y	N	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y	N	Y	Y	Y	N	N	Y	Y	N	N	Y	N	Y	N	18	60.0
0067	Y	N	Y	Y	N	N	N	Y	Y	N	Y	Y	Y	Y	N	Y	N	Y	Y	Y	N	N	Y	N	Y	N	N	N	N	Y	16	53.3
0071	N	Y	N	Y	N	Y	N	Y	Y	Y	N	Y	N	Y	Y	Y	Y	N	N	Y	N	N	N	N	Y	N	Y	N	N	Y	15	50.0
0074	N	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	N	Y	Y	Y	N	Y	N	N	N	N	N	Y	N	N	N	14	46.7
0077	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	N	Y	23	76.7
0079	Y	N	N	N	Y	Y	N	Y	Y	N	Y	Y	Y	Y	N	N	Y	Y	Y	N	N	N	N	N	N	Y	Y	Y	N	Y	16	53.3
0084																															0	0.0
0086	Y	N	N	Y	Y	N	N	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	22	73.3
0091	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	Y	Y	Y	N	N	N	N	N	19	63.3
0094	Y	N	Y	Y	Y	Y	Y	Y	N	N	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	N	Y	Y	N	N	20	66.7
0099																															0	0.0
TOTAL N	11	6	12	7	7	6	14	4	9	18	7	1	9	1	17	5	5	8	1	6	11	19	8	12	7	17	10	12	16	8		

The red cells highlight questions which were answered incorrectly by 10 or more students as discussed in the main body of the text.

Appendix 4: Letter to Headteacher

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

(Name of Head Teacher)
(Name and address of school)

8th September 2021

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 (name of supervisor). In my final research project "Remembering More: Implementing classroom-based 'low stakes' retrieval activities to increase student recall of key knowledge in GCSE Geography", I will explore the ways in which retrieval practice affects students' recall, and whether taking part in retrieval practice affects students' confidence.

The research will take place with my Year 11 Geography class. I intend to develop a programme of retrieval practice in lessons, and routinely quiz students to determine if they are improving in recall of knowledge. My research focus is on how we can challenge students to know and remember more over time, and if students feel they increase in confidence.

By participating in the research, the school would be contributing to a project that will deepen understanding of the utility of retrieval practice in recall of geographical knowledge. There is very little research on retrieval practice for Geography specifically, so this study will contribute to this.

I hope to conduct this research between 04/10/21 and 04/04/21. I would interview students about their feelings towards retrieval practice and invite the whole class to complete short questionnaires at the beginning and end of the study. I will analyse data collected from quizzes completed by the class at half-termly intervals.

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 your formal consent as headteacher is necessary, and not that of individual parents or staff. However, throughout the research, students and other teachers will be able to refuse to participate in any research activities at any time.

All participants, including students, teacher and the school, would be made anonymous in all research reports. The data collected would be kept strictly confidential, available only to my supervisor (and me, and only used for academic purposes. It will be kept for as long as it has academic value.

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. 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,
(Name of Researcher)

Remembering More: Implementing classroom-based 'low stakes' retrieval activities to increase student recall of key knowledge in GCSE Geography.

(Name of Researcher)
University of Oxford, Department of Education

(Name of Head Teacher)

☐ We do not wish to participate in this project.

☐ We would like to find out more about this project.

☐ We would like to take part in this project.

_____
Head teacher's signature

Please return this form to me.

Thank you for your help.