

The effectiveness of recall strategies in the primary classroom to increase pupils' self-confidence in knowing and remembering more

Eleanor Pearson

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	Pearson
First Name	Eleanor
Faculty Board	Education
Title of Dissertation	The effectiveness of recall strategies in the primary classroom to increase pupils' self-confidence in knowing and remembering more.

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

Studies suggest that retrieval of information from long-term memory through quizzing can improve the amount of information learnt over time. Furthermore, performance in quizzes and tests can be linked to pupil self-confidence at school. This study aimed to investigate this effect through the use of ‘quick quizzes’ in a primary school setting, investigating if recall of information through quizzes increases the amount of information learnt and self-confidence in academic ability. The study involved 165 pupils across 4 year groups, with 2 classes per year group. The intervention administered in the study involved each class starting their foundation subject lessons with a quick quiz, for a six-week unit of work. In each year group, one class completed the quizzes in a written format, with the other completing the quizzes orally. Pupil self-confidence was self-reported at the start and end of the unit of work. The results showed that there were no significant differences in the amount of information recalled between the written and oral conditions, however, more pupils from the written condition felt that the quick quizzes contributed to increasing their self-confidence in the topic. Furthermore, teachers found the quizzes to be very successful as an assessment tool in their classrooms. These findings suggest that using recall strategies such as quick quizzes can contribute to a positive classroom environment in which pupils feel self-confident in their learning and that the method of administering these quizzes should be considered for maximum benefit.

Table of Contents

Introduction	5 - 7
Context	5
Literature Review	7 - 37
Long-Term Memory	7
Types of retrieval from the long-term memory	11
Types of recall	12
The importance of recall in the classroom	14
Sticky Learning	16
Curriculum Design	18
Written vs Oral Recall	22
Embedding Recall strategies	26
Pupil Self-Confidence	28
Linking Self-Confidence to Recall	33
Methodology	38 - 54
Summary of Intervention	38
Measures	40
Data Collection	46
Participants	47
Collaboration	48
Ethics	49
Research Limitations	52
Findings	54 - 62
Discussion	62 - 72
Conclusion	72 - 75

References	76 - 89
Appendix A. Teacher Focus Group Questions	90
Appendix B. Self Description Questionnaire	91
Appendix C. Pupil Focus Group Questions	95
Appendix D. Headteacher Letter of Consent	96
Appendix E. SDQ Scores	98

Introduction

The English educational system places a strong emphasis on increasing children and young people's knowledge (Ofsted, 2022). In 2019, school inspectorate Ofsted published findings into curriculum research defining progress as 'knowing more and remembering more' (Ofsted, 2019). This has had significant ramifications for schools across the country as they have deliberated a crucial issue: how can teachers ensure that pupils know and remember more?

Context

In my previous school, there was a drive to adapt the curriculum to meet Ofsted's standards for progress. The school had received a 'Requires Improvement' grading not long before the COVID-19 pandemic began, and the first lockdown in England was used as an intense period of researching and rewriting the curriculum for the new academic year in September. The school focussed on the work of Sherrington (2020) and Rosenshine (2012), embedding recall strategies into the curriculum and hugely improving attainment standards.

After successfully implementing these recall strategies in my previous setting, I have moved to a new school where introducing these is a key part of the School Improvement Plan (SIP). When Ofsted visited in 2019, they listed improvements for the school:

- "Leaders need to continue to review the impact of their new plans for the curriculum when implemented. In particular, they need to check how well the design of the curriculum helps pupils know more and remember more of what they learn across all subjects." (Ofsted, 2019).

- “Leaders should help teachers amend their practice so that lessons help pupils know more and remember more of the curriculum, particularly in subjects other than English and mathematics. Teachers need to make sure they plan opportunities for pupils to revisit old ideas so that they are embedded in pupils’ long-term memory,” (Ofsted, 2019).

These suggestions for my new school meant that I could use this research opportunity to implement the work I did at my previous school and measure the impact that it has on the pupils’ ability to know and remember more.

The recall strategies that were implemented in my previous school were ‘quick quizzes’ (QQs) at the start of each lesson. The QQs would take the format of three questions presented on a slide, asking pupils about content from the lessons they had attended previously. These QQs were used across the entire curriculum and were very repetitive in nature to commit the ‘sticky learning’ to memory. Attainment in core subjects improved over the year, but the impact of the QQs was not measured.

When I moved to my new school, the SIP provided an optimum opportunity for research into the benefits of administering the QQs and the effects that these could have on pupil learning. The school is a state primary in a deprived area of the South West of England, and when I moved here I noticed a distinct lack of pupil self-confidence. Having missed two years of ‘normal’ schooling due to COVID-19, pupils did not seem confident in their academic ability having returned to the classroom. Therefore, pupil self-confidence was another area I wanted to investigate, seeing whether introducing QQs into the curriculum could help to improve pupil self-confidence.

These research ideas and the context of my school led me to develop the following research questions:

1. Are quick quizzes practical and accessible to implement in the primary classroom?

2. How effective are recall strategies in increasing how much pupils know and remember?
3. Do recall strategies increase pupils' self-confidence in their ability to know and remember more?
4. How important is pupil self-confidence in knowing and remembering information?

Literature Review

In this literature review, I will examine current research into recall, apply it to recall in the primary classroom, and then examine recall strategies. I will also investigate the significance of pupils' self-confidence and how this may be related to academic performance. The final section of the review will critically evaluate current research on recall and self-confidence, specifically in the setting of the primary classroom, taking into consideration where pupils' self-confidence comes from and how it can be increased. All of the research cited in the review will be analysed with reference to the questions listed above, demonstrating why this practitioner study is important for understanding pupil self-confidence, as well as for providing guidance for my school as it moves forward with its efforts to give pupils recall opportunities to increase the amount of information they know and remember.

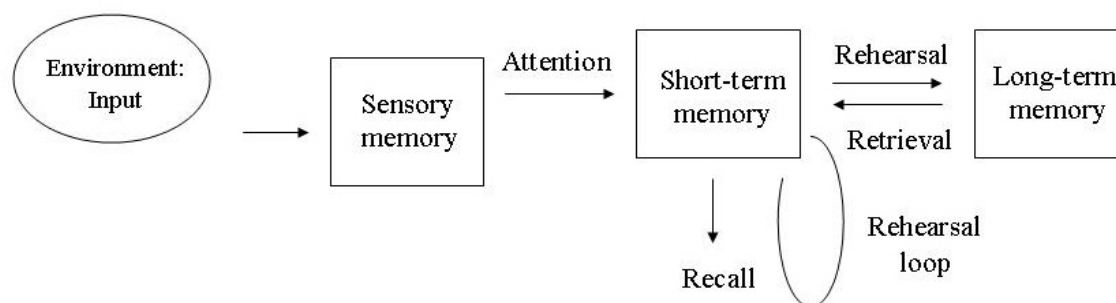
Long-Term Memory

Before considering recall strategies in the primary classroom, it is important to consider why recall is such an important aspect of the acquisition of knowledge; to be able to know and remember more. Research into how the brain works to store knowledge was propelled by Atkinson and Shiffrin (1968) who established the multistore model of memory. In this model there are three separate components: the sensory store, short-term memory (STM) and long-term memory (LTM). Information has to pass through the

memory system in a linear order and can be passed from one component to the other using attention and rehearsal. A diagram of the model can be found in Figure 1.

Figure 1.

Multistore Model of Memory (Atkinson & Shiffrin, 1968)



Information enters the sensory store through the senses, and if attention is applied to this sensory experience, the information will pass into the STM. The STM is 15-30 seconds long (Atkinson and Shiffrin, 1971). For example, if a pupil in a classroom hears a teacher tell them a fact, if they are paying attention, they will remember this fact for 15-30 seconds. The information then passes from the STM to the LTM through rehearsal. Encoding to the LTM is mainly semantic (Baddeley, 2001), although can sometimes be auditory or visual, so if the information is given meaning through elaborative rehearsal, it can be encoded to the LTM. For example, if the pupil thinks about the fact, says it out loud, writes it down, or any other way of rehearsing the information that gives it meaning, it is likely to transfer to the LTM where it can be stored to remember at a later date. This is an active process and information must be retrieved from the LTM regularly so it is not lost (forgotten).

Research is generally supportive of this model of memory. Key studies of memory support that the STM and LTM are two separate stores. For example, numerous studies (see Crano, 1977; Glanzer & Cunitz, 1966; Murdock, 1962) have observed a primacy and recency effect when items in a list are recalled. Items at the beginning of the list are

remembered by the primacy effect, and items at the end of the list are remembered by the recency effect, but items in the middle of the list are forgotten. This demonstrates that the STM and LTM are two separate stores as the information from the start of the list has time to be rehearsed and transferred to the LTM, whereas information at the end of the list is fresh in the memory so still in the STM. The items from the middle of the list are either not rehearsed enough to be transferred to the LTM, or were said longer than 30 seconds ago so have left the STM, so are hence forgotten.

Further support for the multistore model of memory comes from the case of 'KF' (Shallice & Warrington, 1977) who had been in a motorcycle crash and sustained brain damage. KF's LTM was unaffected by the motorcycle crash, but he was not able to transfer new memories to the LTM. He was only able to recall the last bit of information in his STM, which supports the theory that the STM and LTM are two separate stores.

However, there are also limitations to this model. Eysenck (1989) found that the model is oversimplified, for example by suggesting that the STM and LTM are separate, unitary components in which information passes in a linear fashion. For example, there are some types of memories, such as flashbulb memories, where a dramatic event is remembered in great detail, which appear to skip the STM and are encoded straight into the LTM (Conway et al., 1994). Furthermore, it is unlikely that the STM and LTM operate as such simple components. In regards to the LTM, there are many forms of knowledge, such as how to ride a bike, how to calculate a mathematical equation and a conversation you had with someone yesterday, that are unlikely to be stored in the same place (Cowan, 2008). The same can be said for the STM, in which it is unlikely that forms of short-term memory from different cues are all processed and stored in the same way.

Although Atkinson and Shiffrin's multistore model generated a vast amount of research, its shortcomings revealed that there were clear gaps that had not been taken into consideration, particularly with regard to the STM. The working model of memory (Baddeley & Hitch, 1974), another prominent model of memory, aimed to reduce these limitations. According to Baddeley and Hitch, the multistore model offers an overly simplistic image of the STM. The STM is a single component that, according to the multistore model, may store information for a finite period of time with only minimal processing. However, in the working model of memory, the STM has multiple components that can both retain and process information. These components are the central executive, visuo-spatial scratch pad, and phonological loop. A diagram of the model can be found in Figure 2.

Figure 2.

Working Model of Memory (Baddeley & Hitch, 1974)

The figure originally presented here cannot be made freely available via ORA because of copyright. The figure was sourced at Baddeley, A. D., & Hitch, G. (1974). Working memory. In Psychology of learning and motivation (Vol. 8, pp. 47-89). Academic press.

The visuo-spatial sketchpad and phonological loop act as stores, storing and processing visual and auditory information from the senses. The central executive, which functions more like a system, is the most significant part of the working model of memory, allocating information to the visuo-spatial sketchpad, phonological loop and ultimately relaying this information to the LTM. It does this by directing attention to the information and giving priority to the information that is most useful.

In both models of memory, the transfer of information to the LTM is paramount. The LTM has an unlimited duration and capacity (Atkinson & Shiffrin, 1968), therefore information can stay there infinitely. This is the core of learning, if the information is not transferred to the LTM so it can be used at a later date, it is not learnt.

Types of retrieval from the long-term memory

In the multistore model of memory, information is retrieved from the LTM back into the STM where it can be recalled (Atkinson & Shiffrin, 1968). This is an extremely important part of the memory process, as if there is no retrieval of the information in the LTM, there is no point in storing it (Wang et al., 2012). Forgetting happens when the information stored in the LTM cannot be retrieved back to the STM, therefore cannot be recalled (Anderson & Bjork, 1994). There are several theories for why this occurs, for example, the interference theory, retrieval failure and lack of consolidation (see Brown & Lewandowsky, 2010; Ceraso, 1967; Underwood, 1975).

There are several forms of memory retrieval from the LTM, some of which are more applicable to the classroom than others. In all memory retrieval, the neural pathways that were created during the formation of the memory are revisited (Cai, 1990). Four forms of memory retrieval are outlined below:

Recognition. This occurs when information that has been remembered is retrieved through experiencing the memory again (Gillund & Shiffrin, 1984). For example, recognising and remembering someone's name through seeing their picture.

Recollection. Recollection involves rebuilding or piecing together a memory from the LTM again. The brain reconstructs the memory through the use of logic and reasoning, using clues to piece the memory together (Schacter & Buckner, 1998). An example of recollection is remembering the details of a past event, perhaps through looking at pictures as cues to help.

Relearning. This refers to learning information that was learnt in the past but has since been forgotten (Storm et al., 2008). An individual may be able to remember that they have learnt the information before, but not been able to retrieve the information itself. An example could be revising for an exam or assessment; the information has been learnt in a lesson previously but needs to be re-learnt in order for it to be easily retrieved in the exam.

Recall. Recall refers to pulling information from the brain - remembering something without any cues or being in the physical presence of that thing (Gillund & Shiffrin, 1984). Examples include remembering facts, answering questions, or more specifically, remembering a phone number or shopping list. Recall is often used in education, as traditional exams and tests use recall as a way of assessing the amount of information that has been learnt (Browning, 2007).

Types of recall

During memory recall, there is a replaying of the neural activity that took place during encoding (Kragel et al., 2017). When information is encoded to the LTM, neural pathways are formed that link the information to other information already stored, based on meaning. Memory recall involves revisiting these pathways and bringing the

information back to the STM. The speed at which this happens depends on the type of recall, of which there are three:

Free recall. This refers to recall where an individual lists items from their memory, in any order, with no cues to help them. In a psychological study, this may involve listing numbers, objects, words, etc. Free recall has been used to provide evidence for the multistore model of memory, as when free recall is used, a primacy and recency effect can be observed, showing the distinction in nature between the STM and LTM stores (Capitani et al., 1992; Murphy et al., 2006; Steiner & Rain, 1989).

Cued recall. Cued recall refers to recall that takes place when an individual is given cues or prompts, most often semantic. For example, cues could be words, pictures, incomplete sentences, or other prompts based on meaning that allow individuals to recall a specific item. Individuals often remember more items when cues are used than in free recall, as meaning is given to the information and most information is encoded to the LTM this way (Baddeley, 1966).

Serial recall. This refers to recalling items in the specific order that they were presented. This can be used for information stored in either the STM or LTM. For example, recalling a list of digits after they have been presented is a popular way of testing an individual's STM capacity (Schroeder et al., 2012). The capacity of the STM in the average individual was worked out to be approximately 7 items using serial recall in research by Miller (1956). Various studies have shown that the primacy and recency effect can also be observed using serial recall, as items at the start and end of the list are more easily remembered (Crano, 1977). Recall ability decreases when items on the list increase (Miller, 1956).

The importance of recall in the classroom

Retrieval from the LTM is a fundamental practice in the classroom. Research into the 'testing effect' - that active retrieval produces better retention of information than passive reading - has been undertaken for more than a century. Abbott (1909) found that the opportunity for recall by using tests was of benefit for an individual's learning, shortly followed by Gates (1917) and Jones (1923; 1924) who also showed that tests are effective for committing learning to the LTM. The methodology of this research all takes a similar format: a group of participants will learn a set of material, half of the participants will undertake one or more tests on the content of the material, whilst the other half do not take any tests. Participants are then tested on the material sometime later. Numerous studies (see reviews by Agarwal et al., 2010; Roediger & Karpicke, 2006) have shown that participants who were tested on the material remembered more of the content. According to Wheeler and Roediger (1992), taking a test offers a far bigger benefit to the longevity of memory than not taking a test.

However, much of the early research into the testing effect comes from laboratory studies using adult participants, rather than real classrooms with school-aged children. This means that although the testing effect is a real phenomenon in these studies, the results cannot be generalised to classroom settings. A meta-analysis by Schwieren et al. (2017) showed that across 19 publications of the testing effect in psychology classrooms, there was a significant overall effect size, demonstrating that retrieval practice was beneficial to learning outcomes. However, this was a meta-analysis of studies using undergraduate psychology student participants, again adults, who may have some bias towards performing well in these studies. Often undergraduate psychology students are required to participate in studies as part of their course (Nielsen et al., 2017), with Nielsen

et al. explaining that this could bias participants towards performing to support the hypothesis.

To overcome this limitation, researchers have more recently started looking into the impact of testing in real primary and secondary schools. The testing effect is evident and frequent retrieval practice enhances long-term learning in the typical classroom (Agarwal et al., 2009; McDaniel et al., 2011; McDaniel et al., 2012). Agarwal et al. (2012) found that across a five-year period, with more than 1400 middle school pupils, retrieval in real classroom settings improved learning.

Recall is a key form of retrieval that can be used in the classroom to strengthen the LTM (Butler & Roediger, 2007). The strategies that Agarwal et al. (2012) used involved recall quizzes to increase long-term learning. The aim of the study was to apply successful laboratory experiments to real-world classrooms and the effects were significant. In the first part of the experiment, 142 sixth-grade pupils either received three short low-stakes quizzes at various points during the unit of work in addition to normal teaching or received normal teaching with no quizzes. At the end of the unit, quizzed material resulted in significantly better pupil performance than the non-quizzed condition. Pupil performance on quizzed material was 79%, compared to 67% on non-quizzed material, even 1-2 months after the unit had ended. This illustrates the practical advantages of quizzing, and consequently recall.

Similarly, Roediger et al. (2011) undertook three experiments using recall quizzes in the classroom. In experiments one and two, pupils were quizzed on course material three times during the semester using multiple choice questions. Roediger et al. found that material from the quizzes was better recalled during exams than material that had not appeared in the quizzes. Furthermore, pupil performance in the exams improved following quizzing using both free recall and multiple-choice answers. The study

concluded that retrieval practice through quizzing in the classroom can boost academic performance.

Sotola and Crede (2021), who undertook a meta-analysis of the research on the association between frequent low-stakes testing in school classrooms and pupils' academic achievement, provide additional support for this effect. Sotola and Crede identified a moderately significant correlation between the usage of low-stakes quizzes and academic attainment using data from 52 samples. This correlation shows that academic performance increases as quiz usage increases. The researchers argued that because of these strong relationships, that quiz performance was a predictor of later exam performance, therefore low-stakes quizzes contributed to high-stakes exam performance. They also found that using low-stakes quizzes increased a pupil's chance of passing the class. Due to these findings coming from a meta-analysis of 52 samples, this shows that the effects of low-stakes quizzes and the testing effect are clear when applied to real classroom settings. In reference to my school setting, this evidence of low-stakes testing being so effective provides an opportunity for my school to use this strategy to increase the amount that pupils know and remember.

Sticky Learning

There has recently been a drive in promoting 'sticky learning' in teaching. According to Inglis et al. (2014), sticky learning means providing children with knowledge and methods that can be applied across contexts to help them to learn and remember information. In essence, sticky learning is core learning that needs to be remembered in order for a concept to be understood, and for other knowledge to be added to. Understanding how the brain encodes and stores information is fundamental to sticky learning as this learning needs to be retained over time.

Schools, particularly primary, are moving towards a sticky learning model of teaching (Inglis et al., 2014). As Ofsted have defined progress as knowing more and remembering more, they discussed the idea that knowledge is ‘sticky’, meaning the more children know, the more they can learn (Harford, 2018). This demonstrates that cognition and the psychology of inputting and storing information in the LTM are becoming more fundamental to teaching practice so that learning in the classroom is maximised.

Recall strategies play a key part in sticky learning due to the process of retrieval strengthening the information stored in the LTM (Cai, 1990). For example, recalling the key information from a topic each week will mean that more information is remembered at the end of the topic, as shown by the testing effect discussed previously (Agarwal et al., 2009; McDaniel et al., 2011; McDaniel et al., 2012).

The ability to remember this sticky learning means that children will gain more knowledge over time as new knowledge ‘sticks’ to this core learning. The neuroscience behind this is that when the new information is encoded, connections between the neurons are made, so new knowledge ‘sticks’ to the previous information learnt. Furthermore, recalling this information strengthens the connections between these neurons.

This research suggests core sticky learning is fundamental to planning high-quality lessons. In my school, at the start of planning a topic, sticky learning, the information we want the pupils to know at the end of the topic, is put into a single document called a knowledge organiser. This A4 page contains all of the information that the pupils are expected to know, and other additional knowledge can be built upon this. The knowledge organiser forms the basis of the planning for the lessons, to make sure all of the content is covered in a logical and sequential order. Pupils also have access to the knowledge organiser in lessons to aid their information retrieval.

There are arguments for and against the use of knowledge organisers as a way of condensing information that pupils are required to learn. Kirby (2015) explained that knowledge organisers are powerful tools for organising the most vital and useful knowledge that pupils need to learn to be successful in a particular class. When used effectively, knowledge organisers are powerful memory tools and Kirby argues that they can be used to support the testing effect by providing a basis for recall opportunities. However, Kirby also explained that knowledge organisers display very specific knowledge. Hall (2020) argues that a knowledge-led approach detracts away from deep learning, problem-solving, and curriculum subjects where learning is procedural. There are many forms of learning and knowledge that cannot be displayed or organised in this way. Therefore, as well as the use of knowledge organisers in my school to promote sticky learning, there are opportunities in the curriculum and the pedagogy delivered for more recall strategies.

Curriculum Design

The design of the curriculum is an important marker for the quality of education (Harford, 2018). In Ofsted's most recent inspection framework (Ofsted, 2022), schools are judged on the following four categories: quality of education, behaviour and attitudes, personal development, and leadership and management. An overview of these categories can be seen in Figure 3.

Figure 3.*Ofsted Inspection Framework Overview (Ofsted, 2019)*

The figure originally presented here cannot be made freely available via ORA because of copyright. The figure was sourced at Ofsted. (2022). School inspection handbook. GOV.UK. Retrieved 11 May 2022, from <https://www.gov.uk/government/publications/school-inspection-handbook-eif/school-inspection-handbook>

Quality of education can be split into three factors that interlink with each other: intent, implementation and impact. For example, if the intentions of the curriculum are good and it is implemented effectively, the impact will produce good attainment. Curriculum design falls under the ‘intent’ factor and therefore an important aspect in producing a good quality of education.

Research has asked for decades - what makes a high-quality curriculum? The UNESCO International Bureau of Education (IBE) published a document titled this question to support its member states in producing high-quality curriculums to be delivered across the world (Stabback, 2016). The document sets out that education should be (1) inclusive and equitable, (2) characterised by quality learning, (3) promoting lifelong learning, and (4) relevant to holistic development (Stabback, 2016). Furthermore, Stabback proposed that quality learning is “underpinned by a set of theoretical and philosophical beliefs about how children learn” (p. 17).

The research discussed previously has demonstrated that using retrieval strategies is a fundamental component of committing information to the LTM and therefore

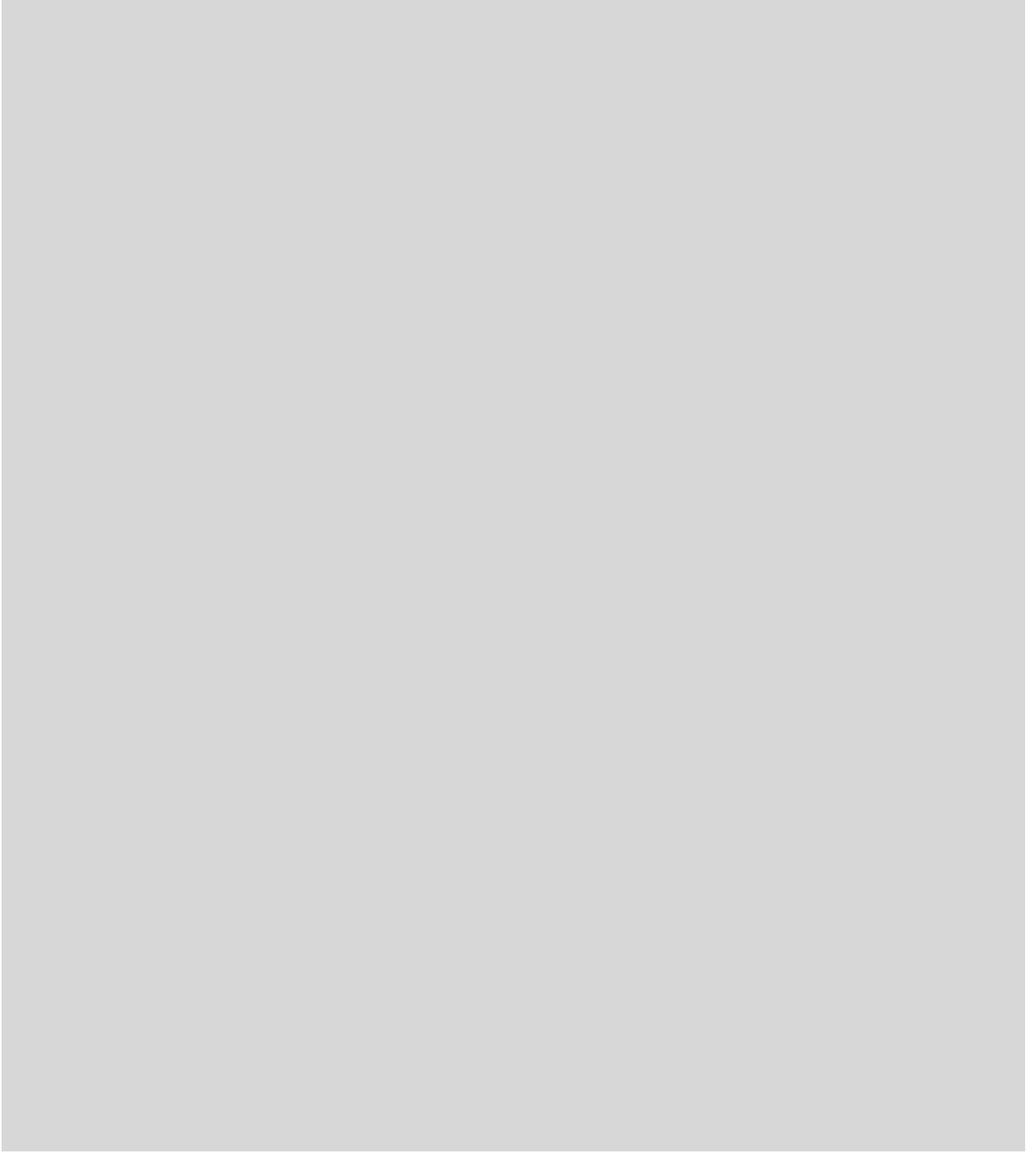
instrumental in how children learn. When working in my previous school setting, the curriculum was completely redesigned and rewritten over the first COVID-19 lockdown period, along with the pedagogy used to implement it, allowing these principles to be put into practice. Much of the research used to inform these curriculum design choices was based on the work by Sherrington (2020) who used Rosenshine's (2012) 'Principles of Instruction' to propose strategies which can make a curriculum meet the benchmarks set out by the IBE.

Sherrington (2020) places an emphasis on retrieval practice in the curriculum. Sherrington simplified the Principles of Instruction to provide an overview of the ten principles that can be easily and successfully incorporated into a school's curriculum in order to achieve maximum impact. These ten principles are found in Figure 4, a poster designed by Caglioli (2016) to summarise the principles.

Figure 4.

Rosenshine's Principles of Instruction (Caglioli, 2016)

The figure originally presented here cannot be made freely available via ORA because of copyright. The figure was sourced at Caglioli, O. (2016). The Principles of Instruction. [Poster]. Retrieved 11 August 2022, from: <https://teachinghow2s.com/news/principles-of-instruction>



As Figure 4 demonstrates, the first principle is ‘daily review’ - a recall technique to retrieve information from the LTM. Furthermore, the tenth principle is ‘weekly and monthly review’, another form of retrieval to embed recently-learned material into the LTM. This demonstrates the need for retrieval strategies, daily, weekly and monthly, to be built into curriculum design to maximise the amount of information that children can learn.

Sherrington (2020) has published several forms of recall strategies that he advises teachers to use in the classroom. Some of these include quick-fire quizzes, silent self-quizzes, paired quizzes, summarising and producing mind-maps. A theme that runs through his recall strategies is that the majority use quizzes, an effective recall technique as shown by previous research discussed in this study (for example Agarwal et al., 2009; Butler & Roediger, 2007; Roediger & Karpicke, 2006). Research has shown that quizzes are an effective tool in the classroom to increase the amount of information learnt over time, and this work from Sherrington demonstrates the ability for this to play a fundamental part in a ‘quality curriculum’ design.

Written vs Oral Recall

The way that recall strategies are delivered is an important consideration. Different types of recall may have different impacts in the classroom. Teachers have to think carefully about the information they are teaching and what would be the most effective way to deliver it to the pupils. For years, a focus of education research was on learning styles and how to cater to the needs of all pupils so that they can best retain the information that is taught to them. Learning styles refer to a range of theories that explain differences in how individuals learn (Coffield et al., 2004), most often categorising individuals as ‘visual learners’, ‘auditory learners’, ‘kinesthetic learners’ etc.

Since 2012, educational research has frequently referred to learning styles as a ‘neuromyth’ (Dekker et al., 2012), often for separating pupils based on their style of learning. Vasquez (2009) argued that defining pupils by their style of learning could create a self-fulfilling prophecy in which pupils develop self-limiting beliefs when trying to learn in a style they do not categorise themselves with.

However, research into learning styles encourages teachers to cater to as many styles as possible within each lesson in order to target every pupil with information in a style they learn in (Romanelli et al., 2009). Although evidence from educational researchers and psychologists has found fundamental flaws in the idea that individuals learn in a specific style and that style only (Rohrer & Pashler, 2012), there are benefits to considering the different designs of inputs even if learning styles are not true. Research from learning style theories has made teachers think about the ways in which they are delivering the content of lessons, considering the preference of learning input for their pupils. For example, the VARK model (Fleming, 2001) proposes that individuals learn in one of four ways: visually, aurally, by reading/writing, or kinaesthetically. Although this model faces criticism, as labelling pupils as one of four styles could be a hindrance to their learning, the general concept of there being different ways of learning through varying types of information input is extremely important for teachers to consider (Fleming, 2012).

Linking back to the models of memory, the working model of memory by Baddeley and Hitch (1974) proposes that the STM is split into two main stores that process and hold information, the visuo-spatial sketchpad and the phonological loop, before the information gets transferred to the LTM. The visuo-spatial sketchpad is responsible for visual and spatial information, whilst the phonological loop is responsible for auditory information, such as spoken or written material. This links to the VARK model as it

proposes that information can be encoded in different ways. Numerous studies show that these two stores are separate as if two tasks that use the same component, such as the phonological loop, are being performed at the same time, then they will not be performed as well as they are separately (Baddeley, 1992). However, when two tasks use the two different components, they should be performed just as well together as they are separately. Studies have shown that where tasks use both the phonological loop and the visuo-spatial sketchpad, retention of information is better (Baddeley & Hitch, 1974).

The way that the recall strategies are delivered, considering the input of information into the memory, is important for the effectiveness of the strategies. Many studies on low-stakes testing and retrieval practice in the classroom (e.g. Agarwal et al., 2001; Sotola & Crede; 2021) have been paper and pencil tests that only cater for a reading/written learning style, which do not suit every individual in the classroom. At present, there are no studies of knowledge that consider the delivery of the recall strategies and how this may alter the outcome of the testing effect.

Some children, particularly SEND pupils, will not be able to access written tests, due to various reasons such as not being able to read the questions, or not being able to write the answers. For example, a child in my current class is extremely knowledgeable about the subjects we are learning, but has a diagnosis of ADHD and gets dysregulated whenever he is asked to complete a writing activity. If he was asked questions about the topic we are studying orally, he would confidently recall the information he had learnt, but if he was asked to sit a written test, even low-stakes, he would inevitably score zero.

Considering this, there is an opportunity to research whether the delivery of the recall strategies or 'style', has an impact on the amount of information that is recalled. Using low-stakes tests with the same layout of two to three questions, as used in previous research discussed in this review, the delivery of the recall strategies could be

investigated. To make the tests accessible for more learners, it would be interesting to investigate whether written paper and pencil recall strategies are actually more beneficial than oral delivery where a child is not required to write any answers down.

This will be interesting to investigate as several points can be raised as to why written recall strategies, although not accessible to everyone, will be more effective for the average child. Written strategies use both the phonological loop and visuo-spatial sketchpad to interpret information, as the individual must read the question, which is presented visually so uses the visuo-spatial sketchpad, but also uses the inner voice to read, which is part of the phonological loop. Additionally, when writing the answer down, both components will be used. Oral quizzes use the phonological loop as the questions are presented aurally and fed back orally. As found previously, tasks which use both components tend to produce better outcomes (Baddeley & Hitch, 1974). However, if recall strategies are presented aurally and are responded to orally, only the phonological loop is being used so outcomes may not be as strong.

Writing information down has been shown to improve the amount of information retained over time (Magen et al., 2015; Schoen, 2012; Smoker et al., 2009; Wiley & Rapp, 2021). Wiley and Rapp found that adults learnt information more successfully when handwriting notes than if they had typed their notes or simply watched videos of the content. They explained that when handwriting, both perception and motor skills are used, therefore the brain is engaging more areas and information is more likely to be learnt. This is supported by Ose Askvik et al. (2020) who found that 12-year-olds and young adults learnt information more efficiently when handwriting notes rather than typing. This study tracked and recorded brain activity and found that handwriting created more neural connections, so that information was stored more effectively. These results indicate that

written recall strategies will be more effective than oral recall strategies, but these are not accessible to all pupils, so is an important question to investigate.

Embedding Recall strategies

There are many recall strategies that can be used to retrieve information, but some are easier to embed into the curriculum than others. In my previous school, the recall strategies by Sherrington were tried and tested over a number of weeks and as a collective, the teachers decided to use low-stakes ‘quick quizzes’ (QQs) for the first principle, ‘daily review’. These QQs were highly successful at increasing the amount of information pupils could recall from one lesson to another.

Also in my previous school, maths, English and reading were taught in the morning, with the remaining curriculum subjects (science, history, geography, art, D&T etc.) each being taught once a week for an hour in the afternoon. Therefore, when using the QQs at the start of each lesson, the children were either retrieving information for a subject they had studied the previous day, or at most in the previous week. Ebbinghaus (1885) found that roughly 66% of newly learned information is forgotten within one day, which explains the necessity of the quizzes to provide a retrieval opportunity so soon after the content is learnt. However, in my new school, the remaining curriculum subjects (for the purpose of this study, referred to as foundation subjects) are taught in unit blocks. For example, in one half term, all of the afternoon lessons (besides physical education) could be geography and in the next half term, all of the afternoon lessons could be history. This means that although information in the QQs would still be recalled from one day to the next, there are also long periods of time where a subject is not taught so information has more chance of being forgotten. Ebbinghaus (1885) found that after six days, approximately 75% of them information learnt is forgotten.

For this reason, and based on the context of the SIP, introducing QQs to my new school was a necessary step. The quantity of literature on retrieval practice and using recall in the classroom is growing and I felt it was important to introduce an intervention that allows me to put this research into practice. The literature, my previous experience and work by Sherrington suggests that QQs are easy and practical to implement into the primary curriculum, however, this study aims to test this in a new context where retrieval practice has not previously been considered and is a new concept to the other teachers collaborating on the study.

The accessibility of the QQs is also an important consideration for me to make. This applies to the accessibility of teachers being able to embed the recall technique into the curriculum, and also the ability of the pupils to access the QQs independently. To embed the recall strategies into the curriculum effectively, research suggests that models for the teachers to use are an effective way to share pedagogy (Maass & Engeln, 2018). As discussed previously, the delivery of the recall strategies is important for accessibility for pupils. Research that has been currently published has not tested whether written or oral QQs are more effective for pupils to know and remember more. Therefore, there is an opportunity for research to explore this.

Embedding the recall strategies into the school's curriculum is not only beneficial for the SIP and the quality of education provided but also helps to close the gap in knowledge pupils have acquired during the COVID-19 pandemic. Kuhfeld et al. (2020) predicted that the effects of COVID-19 school closures will put children behind academically. This gap in attainment will have a knock-on effect on all aspects of the pupils' lives, such as their confidence, self-esteem and motivation.

Pupil Self-Confidence

One of the most significant behavioural motivators in people's daily lives, including in school children, is self-confidence (Bandura, 1986). Numerous studies have been conducted on pupil self-confidence and the factors that contribute to developing it, but few have focused specifically on retrieval practice and the impact of testing. Numerous academics have offered their own views about how learners develop a positive self-concept and gain confidence in their abilities. A rising body of evidence suggests that self-confidence is critical to academic achievement (Bandura, 1977; Ericsson et al., 1993).

A key theory of self-confidence is self-efficacy theory (Bandura, 1977; 1986; Bandura et al., 1999). Self-efficacy refers to a person's confidence in their ability to carry out specific behaviours required for a specific performance. In the context of the testing effect, this means a person's belief of how well they think they will perform in a test and how well this will help them to retain information over time. These self-efficacy beliefs are crucial to many aspects of life, including the selection of goals to strive for, how motivated a person is for a specific goal, and how much energy to put into working towards it.

According to Bandura (1977), an individual's self-efficacy is developed by four main influences: mastery experiences, vicarious experiences, social persuasion, and emotional states.

Mastery experiences. Mastery experiences refer to the experiences that a person gains when they successfully overcome a new challenge. These are said to be the most important influence on self-efficacy as they shape whether a person believes they have what it takes to succeed. Furthermore, not being able to successfully overcome a new challenge undermines self-efficacy as a person is less likely to try new challenges in the

future. In the context of a school, mastery experiences are where pupils overcome a new challenge in the classroom or achieve a goal they have set themselves. In my school, we have ‘sparkle’ certificates that the children are awarded when they reach a goal or have performed particularly well. Self-efficacy theory would assume that these experiences are vital in increasing a pupil’s confidence, as they are more likely to have a go at a new challenge or believe they will achieve more things.

Vicarious experiences. These are experiences of seeing other people successfully complete tasks, and therefore thinking that you would also be able to complete them. If an individual has a positive role model, especially a role model with high self-confidence, they are more likely to have high self-confidence themselves and imitate their role model’s behaviour. Copying the behaviour of role models is a key part of children’s self-development, as theorised in Bandura’s (1969) Social Learning Theory. Role models can include any person an individual looks up to, such as a parent, sibling, grandparent, or teacher. In school, role models are most likely to be peers (Murphey, 1998). If a peer succeeds at a challenge, self-efficacy theory stipulates that a child will be more likely to have a go at that challenge themselves. This is true of my school, where children see others ‘sparkling’ by achieving things, they are more motivated to ‘sparkle’ themselves.

Social persuasion. Social persuasion in self-efficacy theory is the concept that receiving positive feedback whilst undertaking a challenging task will make an individual more likely to believe that they will accomplish the task. This positive feedback increases self-confidence and can then be applied to more challenges in the future. For example, Redmond (2010) stated that self-efficacy is “the influence of encouragement or discouragement pertaining to an individual’s performance or ability to perform”. Giving positive feedback is fundamental to teaching practice (Hattie & Timperley, 2007). Diedrich (2010) found that using positive reinforcement in the classroom through praise

significantly improved pupils' behaviour, motivation and social skills. Self-efficacy theory suggests that praise motivates pupils as they are more likely to be self-confident if others are supporting them and reinforcing the idea that they will successfully complete a challenge. I have certainly observed in my years of teaching that when pupils are praised by myself, other teachers or their peers they become more motivated to complete the task at hand or a new challenge.

Emotional states. The final influence in self-efficacy theory is emotional state. This denotes that the emotional, physiological and psychological well-being of an individual will influence how they feel about their abilities at any certain time. For example, if an individual is struggling with a mental health disorder such as anxiety, they may feel less confident in their ability to perform a certain task. Bandura (1982) theorised that it is much easier to boost self-efficacy when an individual is feeling healthy and well. Alternatively, the level of self-efficacy someone has may also contribute to how these states impact their performance. Someone with a high level of self-efficacy may not let an illness disturb their performance, whilst someone with a low or more fluctuating level of self-efficacy may be more impacted. In the classroom, the emotional and physical state of pupils is fundamental to them feeling happy and confident at school (Cheng & Furnham, 2002). In the context of my school, pupils who require emotional support sessions tend to lack confidence in the classroom and motivation to perform well.

There are numerous other theories of self-confidence. Branden (1969) added to self-efficacy theory by suggesting that self-confidence is made up of two unique components: self-efficacy and self-respect. Branden's definition of self-efficacy is similar to Bandura's, in that he defines it as the confidence an individual has in their ability to overcome life's challenges. This definition is not as detailed as Bandura's, which details the four influences on self-efficacy. Branden defines self-respect as an individual's belief

that they are deserving of success and to overcome these challenges. Branden adds self-respect to self-efficacy because even if an individual knows they are able to overcome challenges, they need to believe that they deserve to.

Both Bandura's (1977; 1986; 1997) and Branden's (1969) theories of self-confidence assume that beliefs about an individual's self-confidence can change over time. This is contradictory to Rosenberg (1965), whose definition of self-confidence uses the assumption that these beliefs are fairly stable over time and difficult to change. Rosenberg is less specific about the factors that influence these beliefs, implying that they are subjective and mean different things to different people, whereas Branden specifies what these beliefs are. Rosenberg created the widely accepted Self-Esteem Scale, which uses 10 statements such as 'I am able to do things as well as most other people' and 'I take a positive attitude toward myself', rated with a Likert scale, to measure an individual's self-esteem. This scale has been used extensively in social science research and is a simple and easy measure to judge an individual's overall self-esteem.

Research indicates that self-confidence is a significant component of academic attainment (Srivastava, 2013). Kleitman and Moscrop (2010) found that in children aged 9-13, higher levels of self-confidence predicted greater academic achievement, regardless of a child's academic ability, gender, age, family dynamics or school fees. Gibson (2008) and Kleitman and Gibson (2011) found a moderate to strong positive correlation between memory, reasoning, self-confidence, and academic self-confidence in primary school pupils. Additionally, a factor measuring self-beliefs explained around 70% of the variance in the measures, indicating that children who have higher self-beliefs in their capacity to reason and remember knowledge also have higher self-beliefs in their academic abilities.

An explanation for higher self-confidence leading to higher academic performance also comes from Gibson (2008) and Kleitman and Gibson (2011), who found a negative

relationship between self-confidence and self-handicapping tendencies. Self-handicapping refers to behaviours such as procrastinating, behaving poorly, turning up late, and other factors that have a negative impact on school performance (Arkin & Oleson, 1998).

Thomas and Gadbois (2007) explained that pupils with low self-confidence often use self-handicapping strategies deliberately, with these strategies being detrimental to learning outcomes, such as poorer exam performance. This is something that I have observed in my teaching practice, particularly during tests and other forms of assessments, as children who lack confidence in their ability will often struggle to apply themselves in test situations or purposely avoid trying so that they don't have to blame poor performance on their ability, but rather a lack of effort.

However, there are limitations to current research into academic self-confidence. Kleitman and Moscrop (2010) found that in children aged 9-13, higher levels of self-confidence predicted greater academic attainment, irrespective of a child's academic ability, gender, etc., however, this greater academic attainment was graded by the pupils' teachers as teacher judgement rather than based on testing providing quantitative scores. Thus, teachers gave higher grades to pupils who self-reported to have high academic self-confidence, with no standardised assessment of academic ability. Kleitman and Moscrop (2010) argued that this means that pupils with higher levels of confidence would also get better school reports, which in turn would positively influence their level of academic self-confidence. This cycle could continue throughout the school years as a self-fulfilling prophecy. This is supported by a longitudinal study by Knapp (2015) which found that low-ability pupils who were graded as low ability at the end of primary school were less likely to finish upper secondary education compared to low-ability pupils who were ungraded, and high ability pupils who were graded as high ability were more likely to finish upper secondary education than high ability pupils who were ungraded. Knapp

theorised that the mechanism for this effect was self-concept, supporting that this cycle can become a self-fulfilling prophecy. However, for there to be a true judgement of whether academic self-confidence positively influences academic performance, there is a need for standardised assessments to be used rather than subjective teacher judgements, as used in these studies. Knapp (2015) used graded and ungraded pupils as participants, but due to the study being completed in Scandinavia, the formal grading was still based on teacher judgements.

Many studies of self-confidence use self-report Likert scales to assess individuals' self-confidence, such as Rosenberg's (1965) Self-Esteem Scale. Kleitman (2008) tested self-confidence using a numerical scale rather than a Likert scale, with 8-9-year-old children rating their academic self-confidence on a scale from 0-100% in reference to how sure their answer to a specific question was correct. This meant the confidence ratings were specific to the academic task at hand and could be compared directly to the performance in the task, rather than relying on a general perception of an individual's academic ability. Kleitman found that these ratings were more accurate measures of self-confidence than the generalised statements used in self-report Likert measures.

Linking Self-Confidence to Recall

The memory model developed by Koriath and Goldsmith (1996) has been used to explain academic self-confidence and its function in academic attainment and recall. Koriath and Goldsmith's approach explains how confidence judgments play a crucial role in memory retrieval, including recall. Retrieval, automatic monitoring, and control are the three phases of the model. Information is retrieved when it is activated from memory, automatically monitored when it is determined whether the information is accurate or not, and controlled when it is decided whether the information should be reported or not.

Using this model, the automatic monitoring and control phases are particularly relevant to academic self-confidence. Kleitman et al. (2012) ascertained that the control phase is vital for academic self-confidence. As the model states that a decision is made on whether or not to report the retrieved information, it can be assumed that different factors will impact this decision. Self-confidence will play a critical part in this as an individual will need to a) be confident that the retrieved information is correct (automatic monitoring) and b) be confident that it is worth reporting that retrieved information (control). Therefore, when an individual is retrieving information, it can be assumed that self-confidence will play a part in this retrieval.

This confidence will be based on the context in which the information is retrieved. Using the model, not only do individuals need to be confident that the information is correct and worth reporting, but also that it is correct in their current social context. In the control phase of the model, individuals can regulate what information to report or not. When tested on the content of a slideshow they had previously seen, 7–12-year-old children were able to increase the accuracy of their answers when given the option to choose which questions to answer, according to research by Koriat et al. (2001), demonstrating that regulation of what information to report is a crucial element of accurate retrieval. An individual may adopt lower criteria for what information to retrieve and report when speaking to a friend than when speaking to a work colleague or employer, or when testifying in court. When a child speaks to a friend or family member rather than a teacher or headteacher, they may also use lower criteria.

This lends itself to low-stakes vs high-stakes testing in the classroom. Much of the work on retrieval practice focuses on the importance of low-stakes testing to build up the amount of information that can be recalled over time. QQs are low-stakes tests as the results do not get recorded towards a pupil's attainment, whereas a high-stakes test such

as an exam would. Cole and Osterlind (2008) found that there were significant differences between test performance when the stakes were higher, demonstrating that the pressures of high-stakes tests impact performance. This is supported by Nicol and Macfarlane-Dick (2006) who claimed that low-stakes assessments are important for practising and generating academic literacy without the fear induced by high-stakes assessments.

Low-stakes testing, such as using recall strategies in the classroom, has been found to improve pupil self-confidence. Meer and Chapman (2014) found that low-stakes assessments designed for improving confidence were excellent for increasing not only the amount of information that pupils could retrieve and retain over time but also for increasing how much the pupils engaged in the unit of work. This piece of practitioner research was conducted over four years at the University of Columbia Business School and used an ‘assessment for confidence’ intervention. Using work by Reid and Barrington (1999), Meer and Chapman theorised that if simple, low-stakes tasks are successful, they will provide pupils with an ‘early win’ that gives them the confidence to complete more complex assessments successfully.

This is supported by research by Simzar et al. (2015) who found that pupils’ confidence and test performance differs by the stakes of the assessment. They found that in maths, self-confidence was more strongly correlated with achievement on the low-stakes assessment than the on high-stakes assessment, showing that confidence is higher when the stakes of the test are lower. Therefore, by using low-stakes quizzes and assessments pupils feel more confident about their achievement in a unit of work or subject.

A theory for the success of low-stakes assessments in increasing pupil self-confidence comes from Sadler’s (1989, 1998) work on formative assessments. Sadler proposed that pupils need a feedback loop in order to benefit from low-stakes

assessments. When pupils have a goal or target to meet, such as performing well in an exam, they need an opportunity for a low-stakes assessment to be able to compare their performance with the goal or target, so that they can engage in an action to close that gap. For example, if pupils have retrieval opportunities at the start of every lesson, as suggested by Sherrington (2020), pupils will know what areas of a topic they are performing well in and what areas they need to work on, so can then approach a high-stakes assessment prepared and more confident.

This lends itself to research by Kenney and Bailey (2021), who found that the use of low-stakes assessments can improve learning over time, but can also reduce overconfidence in university pupils. The low-stakes assessments consisted of a QQ that encouraged pupils to retrieve the information they had learnt in the previous week. At the end of the unit of work, pupils took a high-stakes exam, where the research showed that using QQs increased the amount of information pupils could recall in the exam. Furthermore, pupils were less overconfident on material that had been covered in the low-stakes assessments, but not on material that had not been covered. The researchers concluded that using low-stakes assessments for retrieval practice allows pupils to evaluate their own knowledge of the information being learnt so that they can be more realistic in their expectations for the outcome of a high-stakes assessment.

However, not all research supports this effect. Foster (2021) introduced confidence assessments into low-stakes mathematics assessments for a cohort of 475 pupils across 4 secondary schools, for periods ranging from 3 weeks to a full academic year. Introducing the confidence assessments was to ascertain how confident pupils were that their answers were correct, with pupils rating their confidence between 0 (low) and 10 (high). Foster found that incorporating these assessments, with the aim of increasing pupil self-confidence, did not have an impact on pupil attainment. However, there was also no

detrimental impact, therefore there may not be a link between pupil self-confidence and academic attainment.

Furthermore, to my knowledge, there is currently no literature on low-stakes assessments, such as QQs, and the effect on pupil self-confidence in primary school-age children. Meer and Chapman (2014) found that low-stakes assessments improved pupil self-confidence and attainment, but in university-level students, therefore the findings cannot be applied to school-aged children, as learning and the style of teaching differs hugely between primary and university settings. Moreover, individuals develop and change hugely between primary school, starting at age four, to university where they are at least eighteen years old. Therefore, the results cannot be generalised to a primary setting. Simzar et al.'s (2015) study involved only mathematics lessons, therefore their results are only applicable to mathematics material in those particular settings, and cannot be applied to the entire curriculum.

These mixed outcomes provide further scope for research into the impact of low-stakes assessments, using recall strategies such as QQs, on pupil self-confidence in the classroom. The research on retrieval practice is ever-growing and with new pioneers such as Sherrington (2020) reaching large audiences of teachers to introduce recall strategies into the classroom, this post-COVID-19 period is ideal timing to put these strategies into practice and see if they can have an impact on pupil self-confidence, in an era where pupils are returning to school in-person after time away and are undoubtedly shaken from the events of the past two years. Along with the SIP, this provides a meaningful context in which to conduct this practitioner research and investigate whether implementing these recall strategies into my new school's curriculum will have benefits for both pupil attainment and self-confidence.

Methodology

The study involved implementing a six-week intervention across four year groups in a primary school, in which pupils completed quick quizzes either in a written format or orally, alongside measures of self-confidence. The study had a convergent parallel research design, in which quantitative and qualitative data were produced, before being compared and interpreted together in relation to the research questions.

Summary of Intervention

This study aims to embed current recall strategies in a primary classroom and measure their effectiveness in allowing the pupils to recall more of their knowledge and remember larger quantities of information. To achieve this, both teachers and pupils were involved in the study. The intervention involved implementing QQs across all foundation subject lessons. This was undertaken over a six-week half term, involving eight classes in KS2 (7-11-year-old pupils) at the participating school.

The curriculum at the school is structured so that all afternoon lessons across a particular half term are teaching the same foundation subject, so for this half term, all pupils were studying either a geography or history unit. The QQs were used at the very start of each lesson as an introduction to the lesson, refreshing previous knowledge so that pupils were ready to add to this during the input of the lesson. Each pair of teachers in a particular year group formulated their own QQs based on the learning that they had covered in previous lessons. Teachers were asked to include questions from a variety of lessons, such as a question covering content from the previous lesson, that week, and the previous week. Examples of QQs, as presented to the pupils, can be found in Figure 5.

Figure 5.*Example Quick Quizzes**Quick Quiz*

1. Where is the North Pole: the Arctic or Antarctic?
2. Do you find penguins in the Arctic or Antarctic?
3. What is the capital of England?

Quick Quiz

1. How old was Tutankhamun when he became king, and how old was he when he died?
2. Where was Tutankhamun's tomb discovered?
3. What continent is Egypt in?

As these QQs were built into the lesson slides and had to be planned for by the teachers, both classes in each year group completed them. In one class per year group, the written class, the quizzes were completed individually by pupils using mini-whiteboards and then marked by feeding back with the class. In the other class in the year group, the quizzes were completed orally as a class, discussing the questions as a group, rather than individual children completing all the questions. This decision was made as it allowed all the pupils to be exposed to the questions but also meant that the effect between written and oral answers could be examined. The split of these classes can be found in Table 1. Each class in the participating school has a name, so these names have been changed to preserve anonymity.

Table 1.*Written and Oral Classes*

Written	Oral
Year 3A	Year 3B
Year 4A	Year 4B
Year 5A	Year 5B
Year 6A	Year 6B

Measures

To examine the effects of the intervention, several measures were used to answer the research questions, as presented in Table 2. Detailed explanations of each measure, and how it will be analysed, in the context of the research question it is testing, can be found below.

Table 2.*Summary of the Measures and Analysis for the Research Questions*

Research Question	Measures	Analysis
1. Are quick quizzes practical and accessible to implement in the primary classroom?	- Teacher focus group	- Observation of themes and opinions
2. How effective are recall strategies in increasing how much pupils know and remember?	- End of unit assessment - Pupil focus group	- Scores from the assessments compared across the intervention and control classes using t-tests - Content analysis Observation of themes and opinions

Research Question	Measures	Analysis
3. Do recall strategies increase pupils' self-confidence in their ability to know and remember more?	- Self-Description Questionnaire (Spring, 2004) - Pupil focus group - Teacher focus group	- Differences in the Self-Description Questionnaire at the start and end of the unit compared using t-tests - Content analysis - Observation of themes and opinions
4. How important is pupil self-confidence in knowing and remembering information?	- Self-Description Questionnaire (Spring, 2004) - End of unit assessment	- Correlation analysis between the Self-Description Questionnaire and end of unit assessment

Research question 1: Are quick quizzes practical and accessible to implement in the primary classroom?

This research question was examined using a teacher focus group at the end of the half term to gather teachers' opinions on whether the QQs had been worthwhile and beneficial. A focus group was chosen as this allowed for a mixed-methods approach to be taken for the research. Discussing the outcomes of the QQs with the teachers allowed them to share their opinions on how effective they were and how easy they were to implement in their classrooms.

The questions the teachers were asked can be found in Appendix A. The focus group produced field notes which could then be analysed to answer the research question. Field notes were chosen rather than a full transcript of the focus group as this allowed the researcher to be more present in the discussion, without worrying about transcribing every word. The focus group was not recorded by either audio or video as the aim was to create a safe space for teachers to be authentic and share their opinions. The field notes taken during the focus group were analysed for common themes.

Research question 2: How effective are recall strategies in increasing how much pupils know and remember?

To test this research question, all children across the intervention and control classes were assessed at the end of the unit of work. The assessment took the form of a simple correct/incorrect fact quiz of 10 questions, covering a variety of contents of the QQs over the unit of work. This measure was chosen as the nature of the QQs, being created by the teachers and tailored to the individual units of work, meant that no standardised tests were available or applicable to measure pupil attainment. Before delivering the assessment in each class, the assessments were checked by the researcher to ensure they included a variety of questions to cover the whole unit of work. The same assessment was delivered across the two classes in each year group to ensure comparability of scores.

The assessment produced an average QQ score for each intervention class. T-tests were used to see if there were significant differences in test scores between the written and oral classes. As a qualitative measure, pupil focus groups were also used for this research question, to gauge an opinion from the pupils on whether they felt the QQs had been effective in helping them to know and remember more. An explanation of how pupils were selected for these focus groups can be found under the next research question.

Research question 3: Do recall strategies increase pupils' self-confidence in their ability to know and remember more?

To measure self-confidence in the pupils, several standardised questionnaires were considered. Self-confidence by nature is a subjective concept, so varying instruments and measures have been developed to reliably measure it. For the purpose of this study, the measure needed to be reliable in measuring self-confidence specifically in children, so the

Self-Description Questionnaire (SDQ; Spring, 2004; see Appendix B) was selected. The SDQ is an adapted version of the Self Description Questionnaire-I (Marsh, 1990). Marsh originally developed this questionnaire to measure the different facets of his model of self-concept (Marsh & Shavelson, 1985). This version of the questionnaire was created for children aged 5-12. It measured seven different subscales: Physical Abilities, Physical Appearance, Relations with Peers, Relations with Parents, Reading, Mathematics, and School Subjects. The scales School Subjects, Reading, and Mathematics had 10 items in each scale, whilst the other four scales had 8 items each. Utilizing this questionnaire, several studies have performed factor analyses to validate these subscales and demonstrate the multidimensional nature of self-concept (Marsh et al., 1983).

Additionally, studies have demonstrated the reliability and validity of the questionnaire (Marsh et al., 1984; Marsh & MacDonald Holmes, 1990). For primary and middle school pupils, the SDQ-I was shown to be a valid and reliable self-report measure of motivation, with coefficient alphas for the eight individual scale scores ranging from .80 to .92 (median = .86) indicating good reliability. This demonstrates that, although many decades old, the measure is still appropriate for assessing pupils' levels of self-confidence.

For use with third- and fifth-grade pupils in the Early Childhood Longitudinal Study, the US National Centre for Education Statistics (NCES) condensed the original SDQ-I questionnaire (Spring, 2004). This condensed questionnaire (SDQ) was used in the present study since it only has 42 questions, compared to the SDQ-I's 76, making it simpler to administer to a large group of pupils. In contrast to the SDQ-I's five-point Likert scale, the 42 questions on the SDQ are answered using a four-point Likert scale of "not at all true," "a little bit true," "mainly true," and "extremely true."

In addition, only four of the seven scales from the SDQ-I that were initially published were incorporated in the SDQ. Perceived Interest and Competence in Reading, Math, All Subjects, and Peer Relations were also given new names, and the total number of items for each scale was altered (Tourangeau et al., 2009). There were also two new scales: Internalizing Problems and Externalizing Problems. Through factor analysis, the items for these two new measures were tested and evaluated for validity and reliability (Atkins-Burnett & Meisels, 2001; Pollack et al., 2005). These six scales represented the model of self-concept (Marsh et al., 1988; Shavelson et al., 1976), with the first three scales measuring academic self-concept and the last three scales measuring socioemotional self-concept.

The original scale components were also modified to use wording that was simpler to understand, to make the SDQ questionnaire suited for use with younger children (Pollack et al., 2005). The new items had readability indices using the Flesh-Kinkaid index at a 1.1 grade (Year 2) reading level. To help young pupils make choices, there were fewer possibilities available, with the Likert scale decreasing from five to four.

Only four of the six SDQ scales, Perceived Interest and Competence in Reading, Mathematics, All Subjects, and Peer Relations, were used to answer the study's research objectives, and for this study, they were named Reading, Mathematics, Academic, and Social. Since the items used in the other two scales, Externalizing Problems and Internalizing Problems, were unrelated to the research questions in the current study, they were not included.

Other researchers have utilised the scales in the manner in which NCES modified them (Adelson et al., 2012; Froiland & Oros, 2014; Niehaus & Adelson, 2013). NCES

compared measures of internal consistency of the adapted scales using Cronbach's coefficient alpha to those of the original instrument and found them to be consistent (Ellis et al., 2002).

When administered, the SDQ produced quantitative scores of pupil self-confidence. The pupils completed the SDQ twice, at the start and end of the intervention, and differences in scores were compared using t-tests. This analysis produced a quantitative measure of whether the intervention had caused a significant difference in pupil self-confidence.

As a qualitative supplement to the SDQ, a focus group of four children from each class at the end of the intervention was used, selected based on the scores of their SDQ. The focus groups for each class were made up of the child with the lowest SDQ score in the class at the start of the intervention, the child with the highest SDQ score, and two children with the median SDQ score. This was to ensure a range of self-confidence scores and to determine if their views on the recall strategies and their impacts were similar or dissimilar. Focus groups were chosen as this provided a qualitative data source to compare against the quantitative scores of the SDQ. Data from the pupil focus groups was analysed by looking for common themes using content analysis. The focus groups were not recorded, as this would have required extra parental consent for each pupil, so no transcript was produced. The questions asked to the pupils can be found in Appendix C.

Research question 4: How important is pupil self-confidence in knowing and remembering information?

The SDQ and the assessment scores at the end of the unit of work were used to test this research question. A correlation analysis was conducted to see if there was a significant correlation between self-confidence and score in the assessment.

Data Collection

Table 3 demonstrates how the data was collected across the time frame of the intervention, followed by a detailed explanation of the process.

Table 3.

Data Collection across the Intervention

Week	SDQ	Assessment	Teacher focus group	Pupil focus group
1	✓		✓	
6	✓	✓	✓	✓

Over the course of a six-week period of work that was planned into the school's timetable, data collection was completed. In the first week, teachers gave each pupil a paper copy of the SDQ which they answered by circling the relevant responses. There was no time limit to complete this questionnaire. Before the pupil could hand in their questionnaire, the teachers verified that all of the questions had been completed. Pupils were instructed to complete the questionnaire honestly, without conferring, as the results were being used for research rather than any school-related assessments. All pupils entered their names on the questionnaire so that the researcher could assign ID numbers to match them to their assessment scores, which were stored in a password-protected key document. This was destroyed once all of the data had been logged and matched, thus rendering it anonymous.

In the first week, there was also a teacher focus group to determine predictions about how the teachers thought the intervention would go and to ensure that all teachers were aware of the processes of administering the intervention in their classrooms. All

teachers were involved, with the session lasting approximately 20 minutes. This produced field notes which the researcher kept securely.

At the end of the six-week unit, the SDQ was administered to all pupils for a second time. It was delivered in class, by the class teacher, in the same way as it was delivered in the first week, to ensure consistency. In addition, all pupils completed the end of unit assessment, created by their teacher. The assessment, made up of 10 questions, was delivered as a paper copy, with pupils completing it in pen or pencil. Pupils completed the assessment individually and in standard assessment conditions for the school, sat in their usual seats, but in silence and with no aids to help them. This was to ensure that the pupils' work was their own, but they were informed that the assessment was purely for research purposes and not for formal assessment scores, to relieve any test anxiety.

On the final day of the six-week unit, teacher and pupil focus groups were also held. For the teacher focus groups, teachers in each year group (two teachers at a time) spent 15 minutes with the researcher discussing their opinions on the success of the intervention. As two teachers were in each focus group, there were four focus group sessions in total. The questions used to elicit discussion in the focus groups can be found in Appendix A. For the pupil focus groups, the researcher held eight focus groups of four children each. Focus groups were held in a small room used for group work at the school, so all pupils were familiar with the setting. Focus group sessions lasted 15 minutes. The questions used can be found in Appendix C. Data analysis was conducted when all of the data had been collected.

Participants

A total of 165 pupils ($M = 8.40$ years, $SD = 1.17$, range = 7-11) completed all components of the study. There were 76 female and 89 male participants. 75 extra pupils

started the study but were excluded after missing one or more components of the data collection, including failing to complete all of the questionnaire questions.

Collaboration

There were multiple points of collaboration during the study, both internally in the school where the research was conducted and externally with fellow MSc students. Firstly, I was able to discuss my methodology with other MSc students on course days and this allowed me to fine-tune my methodology and gain useful advice about how to complete my research. When implementing my intervention, nine teachers (including myself) collaborated on the study. After consent was granted and before the study commenced, I spent time with each teacher discussing ideas of how the QQs could be devised and administered. Teachers spent time in their year groups planning the QQs and could decide between themselves which class would be the written condition and which class would be the oral condition.

Near completion of the study, I presented my work at the Poster Conference to the rest of the MSc cohort and using feedback from my peers allowed me to develop the analysis of my findings and think about the implications of my research for my school. When the intervention was complete I presented my findings to my colleagues in the school; those who had collaborated with me on the research and also the KS1 teachers. As a staff body, we decided to roll out the QQs across English and mathematics due to the success of the intervention.

Collaboration in the study was important and allowed me to use others' input to improve my research design and consider how the findings could be used to make a positive impact at my school. Having the other KS2 teachers collaborate with me by carrying out the intervention helped to strengthen staff relationships, particularly as I was a new teacher at the school. However, there were also challenges in collaborating, such as

between teachers who struggle to work together effectively. In one year group, the two teachers both wanted to be the oral condition class, so this was settled fairly using a flip of a coin.

Ethics

This study followed the guidelines set out by the Oxford University Department of Education and The British Educational Research Association's (BERA) Ethical Guidelines for Educational Research (BERA, 2018). As the research involved children, there were many ethical considerations to be made and these are outlined below in relation to five ethical principles for research with children set out by Schenk and Williamson (2005).

Consent and participation. Written consent for the study was granted from the headteacher and all teachers involved in the study (see Appendix D). The headteacher and teachers were informed of the nature of the study in full detail, and no teachers opted not to take part. There were no incentives to take part, other than the positive outcomes that the research could have for the school and SIP. A concern was that teachers would be using a lot of time to plan the QQs and also participate in focus groups, so focus groups were held during specific planning time so teachers did not lose their own time. Teachers were understanding that the research tied in with the SIP therefore planning the quizzes during planning time was of benefit to the school. Pupils were introduced to the research by their teachers and informed that they could withdraw at any time. Both teachers and pupils were informed that they would remain anonymous.

Power relations. Different relations between pupils, teachers and myself as the researcher were acknowledged and efforts to minimise their effects were made. For the teacher focus groups, the teachers knew of the collaborative nature of the research, so

were given time and opportunity to share their opinions on how the research should look and be carried out.

According to Morgan et al. (2002), one of the most important tasks for researchers conducting focus groups with children is to keep an appropriate balance of power in terms of leading and regulating the group and fostering a conversation-friendly environment. Given the inherent power imbalance that results from the researcher being a teacher and thus an authority figure to the participants, this presents a larger challenge to children than to adults. A solution used to minimise this was that the researcher allowed the children to set the scene of the focus group at the start of the session, establishing ground rules for the session which were displayed on flipchart paper throughout. These included ‘everyone gets a chance to speak’, ‘speak one at a time’ and ‘you don’t have to put your hand up to talk’. This gave the children more authority over the session, with the aim that they would be able to share their ideas more authentically.

Furthermore, the size of the focus group was considered. Morgan et al.’s (2002) research into methodological issues of working with focus groups used children aged 7-11-years-old, the same age as the participants in the present study. Morgan et al. found that for focus groups with 7 to 11-year-olds, four to five participants are the most effective. The researchers discovered that with this age group, larger groups made it challenging to promote participatory discussion while also ensuring that the session was not excessively noisy and challenging to record. On the other hand, very small groups of two or three people tended to resemble interviews rather than focus groups, where the researcher spoke to each member in turn and there was less opportunity for participant interaction. Therefore, groups of four pupils were used because it was most practical to implement for this study.

Legal and professional regulations. In addition to following the guidelines set out by the Oxford University Department of Education and BERA, the study was also submitted for approval to the Central University Ethics Committee (CUREC). Data for the study was anonymised and stored in a password-protected document, on the university Nexus 365 encrypted system.

Vulnerable children. The headteacher of the school was presented with a letter (Appendix D) which outlined the plans for the research and consented for all children to take part. Consent from the parents of individual children was not needed as the research fell under ‘Modus Operandi’, allowing me, as a teacher, to conduct research in the classroom or school context. Therefore, all children were able to take part in the study and this was an important consideration for me to make. Faldet and Nes (2021) explained that often vulnerable children are excluded from educational research so they must be included where possible so their voices can be heard. The measures used were devised specifically for use with children so that there would be no psychological harm or risk for vulnerable children to take part. Children with special educational needs (SEN) had their usual adjustments made, such as the use of a scribe, to allow them to complete these measures.

Although the measures used were devised specifically to be used with children, an ethical concern was for the psychological impact on children who did not do well on the QQs as this could damage their self-esteem. When planning the QQs, it was made clear to teachers that of the three questions in the quiz, one or two of the questions would be easy for the children to answer whilst the third was more challenging, so that all children would have a feeling of success in each quiz. Teachers were also instructed to repeat questions across quizzes so that children could recognise questions and recall the answers more easily to build confidence.

Culture and gender considerations. All aspects of this project were designed to carefully take into consideration the pupils' needs. The effects of gender on the SDQ and end of unit results were not investigated as these were not covered by the research questions. All focus groups were mixed gender and included pupils of all cultures and backgrounds.

Research Limitations

There were several limitations to the methodology that should be addressed. Firstly, there was a large number of pupils (75) excluded from the data for not completing all of the components of the study. Pupils were excluded for missing entire components, or for missing out items in a component such as an SDQ question. Many pupils missed at least one component due to isolating with COVID-19, as there was an outbreak at the school during the data collection period. This meant that the study used a much smaller sample than what could have been achieved, rendering the findings less valid as a representation of the entire cohort.

Secondly, I (the researcher) had to self-isolate with COVID-19 during week 6 of the data collection, the week where all four measures were being carried out. The SDQ and QCs were carried out by the other class teachers in my absence, but the teacher and pupil focus groups had to be delayed until I returned after the half-term break. This meant the teachers and pupils had over a week after the QCs ended before they shared their opinion about them.

As the teacher focus groups were delayed, this became problematic as one of the teachers left their post at the school during this week, so they were unable to be interviewed about their thoughts on the QCs. This reduced the teacher focus group sample from nine to eight. Furthermore, on my return to school after recovering from COVID-19, it was discovered that this teacher also had not administered the second SDQ

to their class, so this class (Year 5 - written condition) completed their second SDQ almost two weeks later than the rest of the sample. This had an impact on the results, as discussed in the findings, and affected the validity of the sample, as not all questionnaires were able to be administered at the same time.

Furthermore, the second time that the SDQ was administered (for seven of the eight classes) was during the last week of the data collection period (the last week of the half term) and this also happened to be ‘assessment week’ for the school. During this week the pupils sat reading papers, three maths papers and also completed an assessed piece of writing. Although the pupils were told that the SDQ was not an assessment, pupils were likely feeling more anxious than usual during this week. Kurt et al. (2014) found that assessment periods lead to heightened anxiety in school-aged children. Furthermore, if pupils were feeling anxious, it is likely that their self-confidence would be affected, the very concept that the SDQ was setting out to measure. Coudeville et al. (2011) found that anxiety affects self-confidence. Therefore, it is likely that the SDQ did not produce reliable results the second time it was administered. If the study was repeated, aligning the data collection to be during a ‘typical’ school week rather than a period of assessment would produce more valid results.

Another limitation of the study is that no baseline assessment was conducted at the start of the intervention. Many studies of the testing effect (e.g. Agarwal et al., 2010; McDaniel et al., 2011; McDaniel et al., 2012) have shown a progression in the amount of information remembered over time, for example by doing a baseline assessment at the start of the unit and comparing it to a final assessment at the end of the unit. As collaborators in the study, the teaching staff thought it would be demoralising to test children on content they had not ever seen or been taught before, but that decision meant

that the effectiveness of the QQs as a whole (rather than the written vs oral conditions) cannot be established as there is no baseline measure to compare the QQ scores against.

Findings

1. Are quick quizzes practical and accessible to implement in the primary classroom?

The data from the teacher focus groups found that attitudes towards the QQs were extremely positive. 100% of teachers involved in the study reported that the quizzes were easy or very easy to implement into the curriculum, and 100% said that they felt the QQs had contributed to the children's progress in the unit of work. The QQs were so popular with the teachers for the curriculum units of work that they were then adopted as a starter activity in all English and mathematics lessons too.

Teachers also reported that the QQs were accessible to most children in their class. This was reported to be better for the classes that completed the quizzes orally as all children could engage, whereas in the classes where the quizzes were completed in a written format, not all children could access this. The low-stakes nature of the QQs meant that children who feel anxiety around assessments could engage in the quizzes without feeling pressure from more formal assessments. All teachers said they planned to continue using the quizzes after the intervention had finished.

2. How effective are recall strategies in increasing how much pupils know and remember?

Independent sample t tests were conducted to investigate whether there was a difference in QQ scores between the written and oral conditions in all year groups. The tests were conducted by year group as each year group sat QQs relevant to their year group topics,

therefore the quizzes across year groups were not comparable. The descriptive statistics can be seen in Table 4.

Table 4.

Descriptive Statistics for the Quick Quiz Scores

Year Group	Condition	Number	Mean	Standard Deviation
3	Written	24	5.25	2.13
	Oral	25	5.20	2.33
4	Written	18	6.00	2.77
	Oral	21	4.29	2.59
5	Written	16	7.38	2.31
	Oral	16	7.25	2.70
6	Written	27	7.22	1.93
	Oral	17	7.71	1.86

Shapiro-Wilk's test of normality found that this data was not normally distributed, meaning a non-parametric test was required, so a Mann-Whitney U test was completed for each year group.

Year 3. The test for Year 3 found that there was not a significant difference in QQ score between the written and oral conditions, $U = 281$, $p = .716$, with pupils completing the QQs in the written format ($M = 5.25$, $SD = 2.13$) not scoring significantly higher scores than pupils who completed the QQs orally ($M = 5.20$, $SD = 2.33$).

Year 4. The test for Year 4 found that there was not a significant difference in QQ score between the written and oral conditions, $U = 125$, $p = .073$, with pupils completing the QQs in the written format ($M = 6.00$, $SD = 2.77$) not scoring significantly higher scores than pupils who completed the QQs orally ($M = 4.29$, $SD = 2.59$).

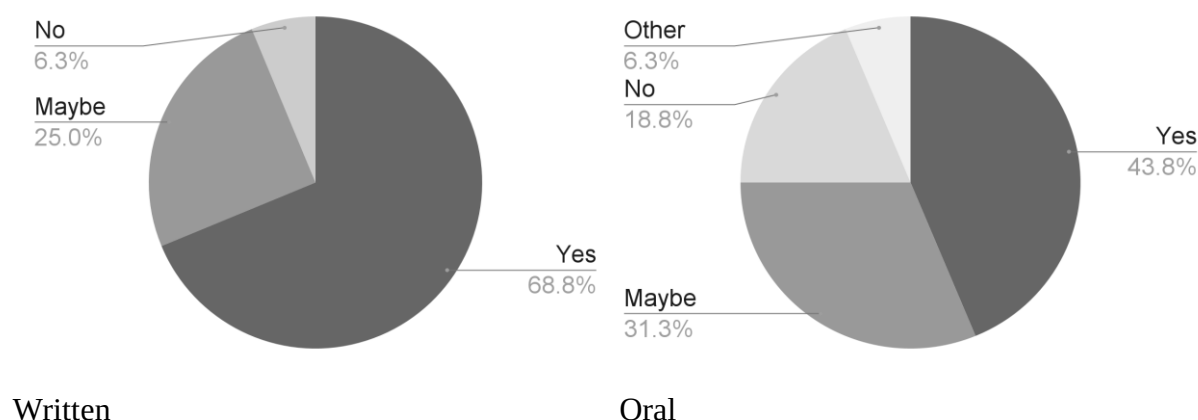
Year 5. The test for Year 5 found that there was not a significant difference in QQ scores between the written and oral conditions, $U = 127$, $p = 1.00$, with pupils completing the QQs in the written format ($M = 7.38$, $SD = 2.31$) not scoring significantly higher scores than the pupils who completed the QQs orally ($M = 7.25$, $SD = 2.70$).

Year 6. The test for Year 6 found that there was not a significant difference in QQ scores between the written and oral conditions, $U = 263$, $p = .416$, with pupils completing the QQs in the written format ($M = 7.22$, $SD = 1.93$) not scoring significantly higher scores than the pupils who completed the QQs orally ($M = 7.71$, $SD = 1.86$).

The results of the t tests show that across all year groups, there were no significant differences in QQ scores between the written and oral conditions. Scores between the two conditions were very similar across all year groups apart from Year 4 where the written condition class scored higher, but not enough to be statistically significant. However, the quantitative data extracted from the pupil focus groups provided a different insight. Of the 16 children interviewed who completed the QQs in the written format, 68.8% reported that they felt the QQs had increased their ability to remember information from the topic, whilst of the 16 children interviewed who completed the QQs orally, only 43.8% reported the same opinion. Figure 6 demonstrates the breakdown of pupil responses to the question ‘Do you think the quick quizzes helped you to remember the information you learnt in the topic?’

Figure 6.

Responses to 'Do you think the quick quizzes helped you to remember the information learnt in the topic?' across the Written and Oral Conditions



The pie charts show that although there were no differences in QQ scores between the written and oral conditions, pupils who completed the QQs in a written format, when interviewed, more often said that the QQs had helped them to remember information on the topic, thus helping their ability to know and remember more.

3. Do recall strategies increase pupils' self-confidence in their ability to know and remember more?

Paired sample t tests were conducted to compare the results of the SDQs at the start and end of the unit of work, to investigate whether using the QQs increased children's self-confidence in the topic. T tests were conducted on the overall SDQ score for each pupil and also for the SDQ subscales: reading, mathematics, academic and social. Shapiro-Wilk's test of normality found that this data was not normally distributed, meaning a non-parametric test was required, so a Wilcoxon Signed-Rank test was completed for each subscale. The results can be found in Table 5.

Table 5.

Paired Sample T Tests for the SDQ and its Subscales from the Start and End of the Unit of Work

SDQ Scale	Variable	Mean	Standard Deviation	Z	P value
Overall	Start	2.735	0.501	1.015	.311
	End	2.701	0.505		
Reading	Start	2.759	0.771	-.570	.568
	End	2.808	0.807		
Mathematics	Start	2.929	0.892	-.415	.679
	End	2.973	0.932		
Academic	Start	2.466	0.707	1.148	.251
	End	2.413	0.737		
Social	Start	2.444	0.682	0.197	.845
	End	2.433	0.749		

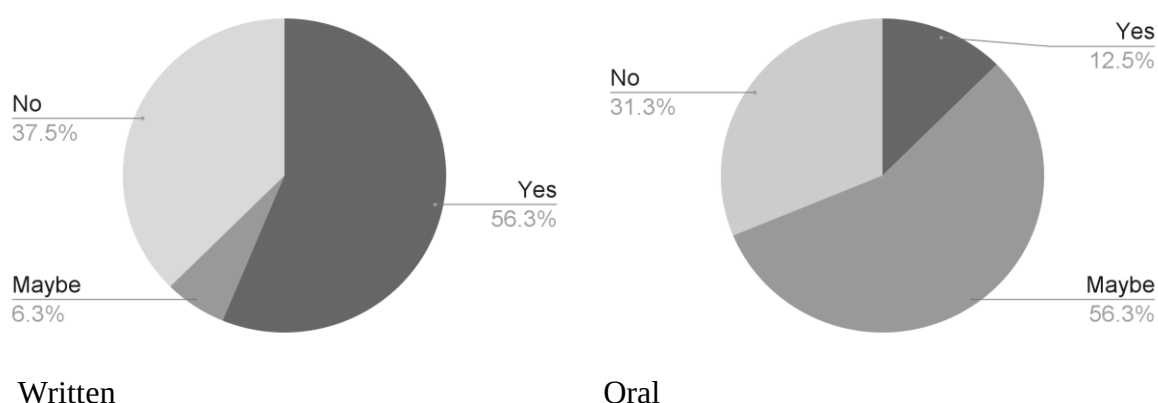
The results from the paired sample t tests show that there were no significant differences in SDQ scores or the subscales between the start and end of the unit of work. In all year groups, the SDQ scores between the start and end of the unit of work remained constant. When looking at the raw data for the SDQ scores (see Appendix E), there was one class, the Year 5 oral condition class, which stood out from the rest of the classes, as SDQ score increased from the start to the end of the unit in three of the subscales: reading, mathematics and academic.

The responses from the 32 pupils interviewed provided similar findings. When asked ‘Do you think the quick quizzes helped your self-confidence in the topic?’ answers were split exactly between ‘yes’ and ‘no’ with 11 children each, with an additional 10 saying ‘maybe’ and one child not being able to remember completing the quizzes.

Considering the responses from the pupil focus groups based on the written and oral conditions provides more insights. Of the 16 pupils interviewed who completed the QQs in the written format, 9 pupils reported that they thought the quizzes helped their self-confidence, compared to only 2 pupils of the 16 who completed the QQs orally. A breakdown of the responses for the written and oral conditions can be seen in Figure 7.

Figure 7.

Responses to ‘Do you think the quick quizzes helped your self-confidence in the topic?’ across the Written and Oral Conditions



The pie charts show that although there were no differences in SDQ scores between the start and end of the unit of work, pupils who completed the QQs in a written format, when interviewed, more often said that the QQs had helped them to be more confident in the topic, thus showing a link between the QQs and self-confidence.

The responses from the teacher focus group support this, with six out of eight teachers reporting that they felt the QQs had helped the self-confidence of the children in their class towards their knowledge of the topic. Of the two that said they did not think the QQs helped, one said that they felt the quizzes ‘possibly’ would have helped, with the effects minimal, whilst the other said that they didn’t think the quizzes helped as they were such a small part of the lesson and information still wasn’t retained across the unit of work.

4. How important is pupil self-confidence in knowing and remembering information?

To test the relationship between QQ score and self-confidence, correlations were completed between the QQ scores and the SDQ scores at the end of the unit of work. Spearman’s rank correlation coefficient was used as Shapiro-Wilk’s test of normality found that again the data was not normally distributed, meaning a non-parametric test was required. Correlations were calculated on the SDQ score as a measure of self-confidence and also the individual subscales of the SDQ, self-confidence in reading, mathematics, academic and social. The correlations can be found in Table 6.

Table 6.*Correlations between the Quick Quiz Scores and Measures of Self-Confidence*

Variable	QQ Score	SDQ Score	Reading	Maths	Academic	Social
QQ Score	-					
SDQ Score	.21**	-				
Reading	.12	.62**	-			
Mathematics	.08	.74**	.32*	-		
Academic	.16*	.77*	.52**	.72**	-	
Social	.08	.51*	.15	.30**	.27**	-

Note. * $p < .05$, ** $p < .01$

The analyses showed that there was a significant correlation between QQ score and SDQ score. This means that pupils who scored highly in the QQs also self-reported to have high self-confidence in the SDQ at the end of the unit of work. There were also significant correlations between the academic subscale of the SDQ with QQ score, but not the other subscales; reading, maths and social. This shows that pupils who self-reported having high academic self-confidence did well in the QQs, but there was no link to QQ score for children that reported high self-confidence in reading, maths or social situations.

Furthermore, the subscales of the SDQ all correlated with the total SDQ score, showing that the SDQ was a valid measure of self-confidence in that the types of self-confidence affected the overall score. Additionally, the reading, mathematics and academic subscales all significantly correlated with each other, showing that self-confidence is universal to the school subjects in these results. The only subscales that did

not have a significant correlation with each other are reading and social, meaning that children who feel confident in reading do not also feel confident in social situations.

As the correlation between the QQ scores and SDQ scores was significant, a simple linear regression analysis was conducted to see whether SDQ score significantly predicted QQ score. The fitted regression model was: $QQ\ score = 6.20 + (3.41 * SDQ\ score)$. The overall regression was statistically significant ($R^2 = .04$, $F(1, 163) = 7.08$, $p < .001$). It was found that SDQ score significantly predicted QQ score ($\beta = 1.03$, $p = .009$).

Discussion

1. Are quick quizzes practical and accessible to implement in the primary classroom?

The results showed that teachers had extremely positive attitudes towards the QQs. This supports previous research by Agarwal et al. (2009) and McDaniel et al. (2011; 2012) that quizzes are a helpful tool for teachers in the classroom. Furthermore, the results showed that all teachers found the QQs easy to implement into the curriculum, again supporting research by Agarwal et al. (2009). When considering how to embed the recall strategies, it was important to acknowledge how easy they would be to implement into the curriculum. Only one of Sherrington's (2020) recall strategies was used, QQs as the first principle 'daily review', but the success puts the rest of the recall strategies in good stead for equal success.

The teachers also all reported that the QQs had contributed to the pupils' progress, supporting research by Roediger and Karpicke (2006) and Roediger et al. (2010) that QQs boost pupil attainment. With years of research into the testing effect, it is in line with predictions that the teachers thought that the quizzes were beneficial to their pupils' progress.

Additionally, the teachers found the QQs so beneficial that not only decided to continue using them after the intervention finished, but they also adopted them in English and mathematics lessons too, supporting research by (Kuo & Hirshman, 1996) that recall in the classroom can be beneficial across the curriculum and is not subject dependent. This demonstrates the success of the intervention as it is being rolled out to a larger number of lessons and also to KS1 as well as KS2.

Furthermore, the results showed that the teachers found the QQs were accessible to most children in their class. This supports research by Gosper (2010) who found that quizzes should be designed in a multidimensional format to be most accessible to participants. This is further supported by the results as the teachers reported that the children who could not engage were children with SEN who could not access the QQs in the written condition. However, in the oral condition, SEN children could participate in the QQs and the low-stakes nature of the quizzes meant that anxiety around assessments was limited.

2. How effective are recall strategies in increasing how much pupils know and remember?

The results showed that in all year groups there were no differences in QQs scores between the written and oral conditions, contrasting research by several researchers (Magen et al., 2015; Ose Askvik et al., 2020; Schoen, 2012; Smoker et al., 2009; Wiley and Rapp, 2021) which suggests that writing information down makes it easier to commit to memory. This demonstrates that in terms of quantitative attainment scores, it does not matter how the QQs are administered, both written and oral have presented the same effect. Looking at the descriptive statistics, the written condition did produce higher QQ scores in three of the four year groups, but this difference was not statistically significant,

therefore the method of administering the quizzes is irrelevant to the overall score and further research is needed to determine if there is an actual effect.

When examining the qualitative data, 68.8% of pupils from the written condition reported that they felt the quizzes had increased their ability to remember information from the topic, whilst only 43.8% of children from the oral condition reported the same opinion. This demonstrates that in terms of opinion, rather than quantitative score, more pupils felt they knew more about the topic, due to the QQs, if they had completed the quizzes in the written format. This is evidence for handwriting being a beneficial tool for memory (Magen et al., 2015; Ose Askvik et al., 2020; Schoen, 2012; Smoker et al., 2009; Wiley & Rapp, 2021), the effect that writing information down by hand commits it to memory better than information that is oral or aural.

Another reason this effect may have been present is that the written quizzes activate more areas of the brain than oral quizzes when they are completed. This supports findings by Ose Askvik et al. (2020) who tracked brain activity when handwriting. Compared to the same task in an oral format, writing answers down produced more activity in different areas of the brain.

Furthermore, referring to the working model of memory, written quizzes could have been more successful for retaining information, in the pupils' opinion, as they involved more components of working memory to be completed. This supports research by Baddeley and Hitch (1974) that tasks which involve multiple components of the working model of memory mean that the information is encoded more successfully. The written quizzes use both the phonological loop and visuo-spatial sketchpad to interpret information, as the pupils must read the question, which is presented visually so uses the visuo-spatial sketchpad, but also uses the inner voice to read, which is part of the

phonological loop. Oral quizzes use the phonological loop as the questions are presented and fed back aurally. Using both components at the same time means that the connections made are stronger and the information is more likely to be remembered. Therefore, it can be said that completing the quizzes in a written format may be a more effective way of helping children to know and remember more.

However, something this study did not have the capacity to investigate was whether the QQs increased attainment over the unit of work, as investigated by Agarwal et al. (2009) McDaniel et al. (2011) and McDaniel et al. (2012). Previous research into recall strategies and the testing effect has shown a progression in the amount of information remembered over time, for example by doing a baseline assessment at the start of the unit and comparing it to a final assessment at the end of the unit. In the present study, there was no baseline score as QQs were not completed at the start of the unit, therefore the effectiveness of the QQs as a whole (rather than the written vs oral conditions) cannot be established.

This is something that may be useful to think about in future research so that clearer findings on the effectiveness of QQs can be found. At present, there have been numerous studies on recall strategies and the testing effect, but there is no literature which compares the methods of carrying out these strategies and whether this has an impact. The present study shows that the pupils felt they knew more in the written condition compared to the oral condition, but there was no quantitative evidence to support this and having no baseline assessment means that comparisons from the start of the unit cannot be drawn. It may have been that the pupils in the written condition had a lower or higher starting point than children in the oral condition but without a baseline assessment this is unknown. If the study was repeated, a baseline assessment could have provided a great opportunity to

test this effect and make it clear whether there is a difference between written and oral delivery of recall strategies.

3. Do recall strategies increase pupils' self-confidence in their ability to know and remember more?

The results showed that there were no significant differences in SDQ scores between the start and end of the unit of work, or for any of the subscales. This shows that the QQs did not change the pupils' self-confidence, using this quantitative measure of self-confidence. This contrasts predictions made based on previous research by Meer and Chapman (2014), who found that the use of low-stakes assessments in the classroom, such as QQs in the present study, can improve pupil self-confidence and also the amount of information that pupils could retrieve and retain over time.

Furthermore, another explanation of why QQs could be expected to raise pupil self-confidence was Sadler's (1989, 1998) feedback loop which proposed that pupils need a feedback loop to benefit from low-stakes assessments such as QQs. When pupils have a goal or target to meet, such as performing well in an exam, they need an opportunity for a low-stakes assessment to be able to compare their performance with the goal or target, so that they can engage in an action to close that gap. However, the findings of the present study show that the QQs did not improve pupil self-confidence, so this feedback loop may not have been present.

There are many reasons, some being limitations to the present study, that this effect has not been observed in the way that the literature predicted. Firstly, the pupils completed the SDQ at the end of the unit of work in the same week as 'assessment week' at the school, a week where the children sit assessments in all core subjects and anxiety is

likely to be increased. Al-Ghwairy and Shuneigat (2006), Bagana et al. (2011) and Kurt et al. (2014) found that assessments lead to increased anxiety and decreased self-confidence. Although the pupils were told at the start of taking the SDQ that it was not an assessment and it did not count for their attainment, the pressures from the other assessments likely led them to lack self-confidence during that week.

This links back to the research discussed in the literature review that performance decreases when the stakes of a test increase (Cole & Osterlind, 2008; Nicol & Macfarlane-Dick, 2006). Cole and Osterlind (2008) concluded that there were significant differences between test performance when the stakes were higher, demonstrating that the pressures of high-stakes tests impact performance. Due to the second SDQ being taken among high-stakes tests, the pupils were under more pressure than usual and this could have had an impact on their self-confidence.

The first time that the pupils sat the SDQ was during a typical school week, with no assessments, where pupil self-confidence could be expected to be average. Due to the second SDQ being taken during assessment week, when stakes were higher, it could be predicted that pupil self-confidence would decrease between the first and second SDQ. However, because there were no significant differences in SDQ scores between the start and end of the unit of work, this result may show that the QQs did help in some way.

In Year 5, the results showed that for the oral condition class, SDQ score increased between the start and the unit of work in three of the subscales. However, in the written condition class, where the teacher left during the study so the second SDQ was not completed until two weeks after assessment week, SDQ score did not increase in any of the subscales. This shows that the SDQ being completed during assessment week could have had an impact on the results, but it is likely that due to the teacher leaving there are

other reasons for a difference in outcome between the two classes. This shows that the limitation caused an anomaly in the data and the results may not be valid for this year group.

Furthermore, the unit of work was only six weeks long, with the SDQ being administered in weeks one and six, so only a gap of five weeks. This is a very short time to expect an increase in scores when looking at the nature of the questions in the SDQ. Five weeks is a short time for a pupil to change their opinion on statements such as 'I like maths' and 'I like reading', especially when there were only 4 Likert scale options. Moreover, the SDQ is a generalised questionnaire about the core aspects of academia and is not specific in any way to the QQs. It may have been better to create a specific questionnaire to judge pupil self-confidence in relation to the QQs and the topics being studied by the children, but that was beyond the scope of this study. The SDQ assessed whether children feel confident in reading, mathematics, generally in academic subjects and social relationships, but all the pupils completed QQs in geography or history, subjects that were not specifically covered by the SDQ. If this study was to be repeated a more specific questionnaire could be created to assess pupil self-confidence in the specific subjects being assessed by the QQs, and also it should be administered over a longer period than five weeks so that there is time for changes to take effect.

Although the quantitative data did not show an improvement in pupil self-confidence, the qualitative data paints a different picture. The results showed that three-quarters of the teachers thought that the quizzes improved pupil self-confidence in the topic, however, the responses of the 32 pupils interviewed showed mixed responses. When split between the written and oral conditions, the results supported that using written quizzes is more effective in improving self-confidence than using oral quizzes.

The quantitative and qualitative data from the pupil interviews, overall, produced the same results. There was no difference in SDQ scores across all pupils from the start and end of the unit of work, and in the pupil interviews, the average response was that there had been no change in confidence. However, when split into the written and oral conditions, the written condition was much more favourable at increasing pupil self-confidence, which means that writing the answers down could have influenced pupil self-confidence. As mentioned previously, it could have been expected that the SDQ score would decrease with the second SDQ being administered during a high-stakes week, but the non-significance of the results shows that the QQs could have helped with keeping pupil self-confidence constant.

The difference between the results from the written and oral conditions for the pupil interviews could be explained by a stronger feedback loop (Sadler, 1989; 1998) being created in the written condition. In the written condition, the questions were presented visually, the answers were written down on mini whiteboards, and then the answers were discussed orally. When the answers were discussed, the pupils were able to tick or cross their answers on their mini whiteboards. By doing this, they were able to see their performance in the QQ as right or wrong answers in front of them. In the oral condition, the pupils discussed the answers together before feeding back as a class and they had no ticks or crosses to represent whether they were correct or not, just the knowledge in their heads. To create a feedback loop to build confidence, an individual must have a goal or target, be able to compare their performance against the goal or target, and then be able to adjust their performance to meet the goal or target (Sadler, 1989; 1998). In the written condition, children were more able to compare their performance against the target as they could see correct and incorrect answers in front of them. Seeing these correct

answers could have increased their confidence, and any incorrect answers meant they knew exactly what to work on to do better in the next QQ, again meaning they would gain confidence.

Furthermore, as the assessment at the end of the unit of work to work out how effective the QQs were was a written assessment, the pupils in the written condition were more acquainted with writing down their QQ answers than the children in the oral condition, who had never had to write down their answers before, only discuss their answers in a small group. This could have led to the pupils in the oral condition feeling less confident, as they had to complete the end of unit assessment independently and were not able to discuss their answers as they were used to.

The qualitative data also showed that three-quarters of the teachers thought that the quizzes improved pupil self-confidence in the topic, supporting research by Meer and Chapman (2014) that low-stakes assessments are excellent for improving confidence and the amount of information pupils can retain. This is a beneficial finding as pupil self-confidence is important for not only academic attainment (Francis et al., 2017) but for overall quality of life in school-aged children (Marriage & Cummins, 2004). However, one-quarter of teachers thought that the quizzes did not help the pupils to feel more confident in the topics they were studying, and this leads to questions that could be investigated in future research: why do quizzes not improve self-confidence for all pupils? How can the quizzes be improved in a manner that would increase pupil self-confidence? These questions provide scope for future research opportunities so that recall strategies can be best implemented to help all children in the classroom to feel more self-confident.

4. How important is pupil self-confidence in knowing and remembering information?

The results showed that QQ score and SDQ score were significantly correlated, showing that the better the academic performance, the more self-confidence a pupil has. This supports research by Gibon (2008) and Kleitman and Gibson (2011), who found that in primary school children, there were moderate to strong positive correlations between memory, reasoning, self-confidence and academic self-confidence. Furthermore, results showed that of the subscales of the SDQ, only academic self-confidence was correlated with QQ score. This shows that pupils who were confident about their overall academic ability also scored highly in the QQ assessment.

The other subscales of the SDQ - reading, mathematics and social, did not significantly correlate with QQ score. This could strengthen the evidence that the SDQ is a valid measure of self-confidence, as the only subscale that the SDQ that relates directly to the QQs is academic self-confidence, the subscale that did correlate. This is due to the QQ assessment covering content from either a history or geography unit, not reading, mathematics, or a pupil's social confidence. Had the QQs been administered across the entire primary curriculum during the intervention, as they have been since due to the intervention's success, there may have also been correlations between QQ score and also the reading and mathematics subscales.

Due to the SDQ and QQ scores being significantly correlated, further analysis was conducted which showed that SDQ score also significantly predicted QQ score. This means that if a pupil had a high SDQ score, this could significantly predict that they also have a high QQ score. This supports previous research by Kleitman and Moscrop (2010) who found that higher levels of self-confidence predicted greater academic attainment, irrespective of a child's academic ability, gender, age, family dynamics or school fees,

and research by Strivistava (2013) who found that self-confidence is a key factor in academic attainment.

Kleitman and Moscrop (2010) argued that this means pupils with higher levels of confidence would also get better school reports, which in turn would positively influence their level of academic self-confidence. This links back to the feedback loop (Sadler, 1989; 1998), where an individual's self-confidence becomes a self-fulfilling prophecy. If a pupil does well in a QQ, their self-confidence will be positively influenced, which over a period of repeated QQs, will mean that their academic self-confidence is increased. Furthermore, if the pupil did not perform well in a QQ, they could use this information to adjust their performance so that the next QQ is improved, and this level of self-awareness of performance is paramount for self-confidence.

Conclusion

In conclusion, the study showed that QQs are an effective recall strategy to use in the primary classroom to increase pupils' self-confidence in knowing and remembering more. The findings produced by the study can be used to answer the research questions set out at the beginning of the study and provide some useful knowledge moving forwards.

Firstly, the results showed that QQs are practical and accessible to implement in the primary classroom, with all of the teachers supporting the introduction of the quizzes and rolling them out to the entire curriculum as a result of the success of the intervention. Next, the results showed that teachers and pupils perceive the QQs to be effective in increasing how much pupils know and remember, although this wasn't supported by the quantitative data. Pupils who completed the QQs in a written format were more positive about how effective the QQs were than pupils who completed the QQs orally. Similarly,

the study showed that pupils perceived the QQs to increase their self-confidence in the topic studied, although again this effect was not demonstrated by the quantitative measure used. Lastly, the study showed that self-confidence is a significant predictor of a pupil being able to know and remember information. This supports and adds to current literature on academic self-confidence and the use of recall strategies.

These conclusions provide many implications for educational research and my practice as a primary teacher and collaborative researcher. This study has added to current literature on recall strategies, low-stakes assessments and the testing effect by investigating the effects of the delivery of the QQs and showing that there is a difference in impact between written and oral quizzes. Furthermore, the study adds to research on the effect of handwriting over oral tasks to show that using handwriting means that pupils perceive that they know and remember more information. Finally, the study supports current research that self-confidence is a predictor of academic ability, and adds to this by finding that completing QQs in a written format increases confidence more than orally, something that has not been investigated before in this context.

For my school, these findings are instrumental. As the intervention was successful, with the teachers collaborating on the study finding the QQs so effective, the QQs were rolled out across the curriculum at the end of the intervention and have been extremely useful in helping children to feel more confident in their learning. At the time of writing, the first full school year has been completed without any disruption from COVID-19. Current school data shows that the pupils have up to two years' worth of gaps in their learning, and using recall techniques has the potential to accelerate progress to help close these gaps. Additionally, the finding that the pupils perceive their confidence to have increased, particularly in the written condition, is extremely important after two years of disruption where confidence was inevitably knocked.

The study has also been very important for my professional development as an early career teacher and for my relationships with my colleagues. Coming to a new school amid COVID-19 recovery and immediately asking colleagues to collaborate in a project of this scale, where they would have to carry out the intervention alongside me, helped to strengthen my relationship with my colleagues immediately and allowed me to use my research to help to work towards the SIP from the onset. The findings from the study help towards the SIP, where Ofsted (2019) had noted that there were not enough recall opportunities as pupils go a long time between units of work, and prepare the school well for this not to be a criticism in the future. Over time, the benefits of the QQs being rolled out across the curriculum will be even clearer and pupils will have a much more secure knowledge of their topics - days, weeks and months after they are taught.

The study also provides many opportunities for future research, be it my own research in the school in the future or more generally in education. If this study were to be repeated, a baseline measure of attainment at the start of the unit of work would have been extremely useful, to gauge quantitatively how much the QQs had increased the pupils' ability to know and remember more. However, the study did show the effectiveness of the QQs in many aspects, and their basis on Sherrington's (2020) work puts the rest of Sherrington's strategies in good stead for also impacting pupil attainment. Future research could investigate more of these strategies and perhaps compare how effective they are against each other. Furthermore, this study tested written vs oral low-stakes assessments made up of three questions each, but there are many other forms of low-stakes assessments and many ways to deliver them that could also be investigated.

As teachers and educational researchers, the most fundamental part of teaching and researching is to help children to achieve their full potential. Using the QQs as an intervention was a method to help children to know and remember more of the topic they

were studying, and these were found to be effective, but there are many other ways that could also be investigated. The study shows that Sherrington (2020) has provided an effective strategy and there are many others he suggests which may be more or less effective. In the future, I may collaborate with my colleagues again to introduce more of these strategies to help close these gaps from COVID-19. However, the findings from the present study have been instrumental for my school, having been rolled out across the curriculum, and hopefully, the effects of introducing this intervention will be felt by pupils for years to come.

References

- Abott, E. E. (1909). On the analysis of the factor of recall in the learning process. *The Psychological Review: Monograph Supplements*, 11(1), 159.
- Adelson, J. L., McCoach, D. B., & Gavin, M. K. (2012). Examining the effects of gifted programming in mathematics and reading using the ECLS-K. *Gifted Child Quarterly*, 56(1), 25-39.
- Agarwal, P.K., Roediger, H.L., McDaniel, M.A., McDermott, K.B. (2009). Feedback increases middle school students' resolution and retention of correct responses. Poster presented at the 50th Annual Meeting of the Psychonomic Society, Boston.
- Agarwal, P. K., Roediger III, H. L., McDaniel, M. A., & McDermott, K. B. (2010). Improving Student Learning through the Use of Classroom Quizzes: Three Years of Evidence from the Columbia Middle School Project. *Society for Research on Educational Effectiveness*.
- Al-Ghwairi, A., & Shuneigat, W. (2006). The impact of exams on anxiety levels among university students. *JRMS*, 13(2), 14-19.
- Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long-term memory: A new taxonomy.
- Arkin, R. M., & Oleson, K. C. (1998). Self-handicapping.
- Atkins-Burnett, S., & Meisels, S. J. (2001). Measures of Socio-Emotional Development in Middle Childhood. Working Paper No. 2001
- Atkinson, R. C., & Shiffrin, R. M. (1968). Chapter: Human memory: A proposed system and its control processes. In Spence, K. W., & Spence, J. T. *The psychology of learning and motivation* (Volume 2). New York: Academic Press. pp. 89–195.
- Atkinson, R. C., & Shiffrin, R. M. (1971). *The control processes of short-term memory*. Institute for Mathematical Studies in the Social Sciences, Stanford University.

- Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *The Quarterly journal of experimental psychology*, 18(4), 302-309.
- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556-559.
- Baddeley, A. (2001). The concept of episodic memory. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1413), 1345-1350.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.
- Bagana, E., Raci, A., & Lupu, L. (2011). Self esteem, optimism and exams' anxiety among high school students. *Procedia-Social and Behavioral Sciences*, 30, 1331-1338.
- Bandura, A. (1969). Social-learning theory of identificatory processes. *Handbook of socialization theory and research*, 213, 262.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American psychologist*, 37(2), 122.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A., Freeman, W. H., & Lightsey, R. (1999). Self-efficacy: The exercise of control.
- Branden, N. (1969). *The Psychology of Self Esteem*. New York: BantamBooks.
- Brown, G. D., & Lewandowsky, S. (2010). Forgetting in memory models: Arguments against trace decay and consolidation failure. *Forgetting*, 49-75.

- Browning, W. G. (2007). *Cliffs Memory Power for Exams*. Houghton Mifflin Harcourt.
- Butler, A. C., & Roediger III, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, 19(4-5), 514-527.
- Caglioli, O. (2016). *The Principles of Instruction*. [Poster]. Retrieved 11 August 2022, from: <https://teachinghow2s.com/news/principles-of-instruction>
- Cai, Z. J. (1990). The neural mechanism of declarative memory consolidation and retrieval: A hypothesis. *Neuroscience & Biobehavioral Reviews*, 14(3), 295-304.
- Capitani, E., Della Sala, S., Logie, R. H., & Spinnler, H. (1992). Recency, primacy, and memory: Reappraising and standardising the serial position curve. *Cortex*, 28(3), 315-342.
- Ceraso, J. (1967). The interference theory of forgetting. *Scientific American*, 217(4), 117-127.
- Cheng, H., & Furnham, A. (2002). Personality, peer relations, and self-confidence as predictors of happiness and loneliness. *Journal of adolescence*, 25(3), 327-339.
- Coffield, F., Moseley, D., Hall, E., Ecclestone, K., Coffield, F., Moseley, D., & Ecclestone, K. (2004). Learning styles and pedagogy in post-16 learning: A systematic and critical review.
- Cole, J. S., & Osterlind, S. J. (2008). Investigating differences between low-and high-stakes test performance on a general education exam. *The Journal of General Education*, 57(2), 119-130.
- Conway, M. A., Anderson, S. J., Larsen, S. F., Donnelly, C. M., McDaniel, M. A., McClelland, A. G., & Logie, R. H. (1994). The formation of flashbulb memories. *Memory & Cognition*, 22(3), 326-343.

- Coudevylle, G. R., Gernigon, C., & Ginis, K. A. M. (2011). Self-esteem, self-confidence, anxiety and claimed self-handicapping: A mediational analysis. *Psychology of Sport and Exercise*, 12(6), 670-675.
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory?. *Progress in brain research*, 169, 323-338.
- Crano, W. D. (1977). Primacy versus recency in retention of information and opinion change. *The Journal of Social Psychology*, 101(1), 87-96.
- Dekker, S., Lee, N. C., Howard-Jones, P., & Jolles, J. (2012). Neuromyths in education: Prevalence and predictors of misconceptions among teachers. *Frontiers in psychology*, 429.
- Diedrich, J. L. (2010). *Motivating students using positive reinforcement* (Doctoral dissertation).
- Ebbinghaus, H. (1885/1913). *Memory: A Contribution to Experimental Psychology*. (Transl: H. A. Ruger & C. E. Bussenius.) New York: Teachers College.
- Ellis, L. A., Marsh, H. W., & Richards, G. E. (2002). A brief version of the Self-Description Questionnaire II. In *Proceedings of the 2nd International Biennial Conference Self-Concept Research: Driving International Research Agenda*. Sydney: University of Western Sydney, SELF Research Centre.
- Ericsson, K.A., Krampe, R.T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review* 100 (3), 363.
- Eysenck, M. W. (1989). Human learning. In *Human and machine problem solving* (pp. 289-315). Springer, Boston, MA.
- Faldet, A. C., & Nes, K. (2021). Valuing vulnerable children's voices in educational research. *International Journal of Inclusive Education*, 1-16.

- Fleming, N. D. (2001). *Teaching and learning styles: VARK strategies*. Neil Fleming.
- Fleming, N. D. (2012). The case against learning styles: “There is no evidence...”.
Obtenido de Designer of the VARK Questionnaire for Learners' preferences,
<http://vark-learn.com/introduction-to-vark/articles>.
- Foster, C. (2021). Implementing confidence assessment in low-stakes, formative mathematics assessments. *International Journal of Science and Mathematics Education*, 1-19.
- Francis, B., Connolly, P., Archer, L., Hodgen, J., Mazenod, A., Pepper, D., ... & Travers, M. C. (2017). Attainment Grouping as self-fulfilling prophesy? A mixed methods exploration of self confidence and set level among Year 7 students. *International Journal of Educational Research*, 86, 96-108.
- Froiland, J. M., & Oros, E. (2014). Intrinsic motivation, perceived competence and classroom engagement as longitudinal predictors of adolescent reading achievement. *Educational Psychology*, 34(2), 119-132.
- Gates, A. I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, 40, 104.
- Gibson, J. (2008). The effects of the learning environment on the metacognitive beliefs, selfconfidence and use of self-handicapping strategies of sixth grade students. *Unpublished Hon. thesis, The University of Sydney, Sydney, Australia*.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological review*, 91(1), 1.
- Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of oral learning and oral behavior*, 5(4), 351-360.
- Gosper, M. (2010). Designing online quizzes: A whole of curriculum approach. In *IADIS International Conference on Cognition and Exploratory Learning in the*

Digital Age, CELDA 2010 (pp. 11-18). International Association for Development of the Information Society (IADIS).

Hall, S. (2020). Knowledge Organisers – A Failed Revolution. *S Hall Teach*.

<https://shallteach.wordpress.com/2020/07/25/knowledge-organisers-a-failed-revolution/>

Harford, S. (2018). School inspection update: education inspection framework January

2010. Retrieved 17 July 2022, from:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/772056/School_inspection_update_-_January_2019_Special_Edition_180119.pdf

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.

Inglis, H. J., Dawson, K. L., & Nishioka, R. Y. (2014). *Sticky learning: How neuroscience supports teaching that's remembered*. Augsburg Fortress Publishers.

Jones, H. E. (1923). *Experimental studies of college teaching: The effect of examination on permanence of learning*. Columbia University.

Jones, H. E. (1924). Behaviorism and Psychology. *The Journal of Philosophy*, 21(7), 193-194.

Kenney, K. L., & Bailey, H. (2021). Low-stakes quizzes improve learning and reduce overconfidence in college students. *Journal of the Scholarship of Teaching and Learning*, 21(2).

Kirby, K. (2015). Knowledge Organisers. *Pragmatic Reform*. Retrieved 11 August 2022, from: <https://pragmaticreform.wordpress.com/2015/03/28/knowledge-organisers/>

- Klapp, A. (2015). Does grading affect educational attainment? A longitudinal study. *Assessment in Education: Principles, Policy & Practice*, 22(3), 302-323.
- Kleitman, S. (2008). *Metacognition in the rationality debate: Self-confidence and its calibration*. VDM Publishing.
- Kleitman, S., & Gibson, J. (2011). Metacognitive beliefs, self-confidence and primary learning environment of sixth grade students. *Learning and Individual Differences*, 21(6), 728-735.
- Kleitman, S., & Moscrop, T. (2010). Self-confidence and academic achievements in primary-school children: Their relationships and links to parental bonds, intelligence, age, and gender. In *Trends and prospects in metacognition research* (pp. 293-326). Springer, Boston, MA.
- Kleitman, S., Stankov, L., Allwood, C. M., Young, S., & Mak, K. K. L. (2012). Metacognitive self-confidence in school-aged children. In *Self-directed learning oriented assessments in the Asia-Pacific* (pp. 139-153). Springer, Dordrecht.
- Koriat, A., & Goldsmith, M. (1996). Memory metaphors and the real-life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences*, 19(2), 167-188.
- Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children's testimony: Can children control the accuracy of their memory reports?. *Journal of Experimental Child Psychology*, 79(4), 405-437.
- Kragel, J. E., Ezzyat, Y., Sperling, M. R., Gorniak, R., Worrell, G. A., Berry, B. M., ... & Kahana, M. J. (2017). Similar patterns of neural activity predict memory function during encoding and retrieval. *Neuroimage*, 155, 60-71.
- Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the Potential Impacts of COVID-19 School Closures on Academic

- Achievement. EdWorkingPaper No. 20-226. *Annenberg Institute for School Reform at Brown University*.
- Kuo, T. M., & Hirshman, E. (1996). Investigations of the testing effect. *The American Journal of Psychology*, 451-464.
- Kurt, A. S., Balci, S., & Kose, D. (2014). Test anxiety levels and related factors: Students preparing for university exams. *J Pak Med Assoc*, 64(11), 1235-9.
- Maass, K., & Engeln, K. (2018). Impact of professional development involving modelling on teachers and their teaching. *ZDM*, 50(1), 273-285.
- Mangen, A., Anda, L. G., Oxborough, G. H., & Brännick, K. (2015). Handwriting versus keyboard writing: effect on word recall. *Journal of writing research*, 7(2), 227-247.
- Marriage, K., & Cummins, R. A. (2004). Subjective quality of life and self-esteem in children: The role of primary and secondary control in coping with everyday stress. In *Quality-of-Life Research on Children and Adolescents* (pp. 107-122). Springer, Dordrecht.
- Marsh, H. W., Barnes, J., Cairns, L., & Tidman, M. (1984). Self-Description Questionnaire: Age and sex effects in the structure and level of self-concept for preadolescent children. *Journal of Educational psychology*, 76(5), 940.
- Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of educational psychology*, 80(3), 366.
- Marsh, H. W., & MacDonald Holmes, I. W. (1990). Multidimensional self-concepts: Construct validation of responses by children. *American Educational Research Journal*, 27(1), 89-117.

- Marsh, H. W., Relich, J. D., & Smith, I. D. (1983). Self-concept: The construct validity of interpretations based upon the SDQ. *Journal of Personality and social psychology*, 45(1), 173.
- Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational psychologist*, 20(3), 107-123.
- McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger III, H. L. (2011). Test-enhanced learning in a middle school science classroom: the effects of quiz frequency and placement. *Journal of Educational Psychology*, 103(2), 399.
- McDaniel, M. A., Thomas, R. C., Agarwal, P. K., McDermott, K. B., & Roediger, H. L. (2012). Quizzing promotes transfer of target principles in middle school science: Benefits on classroom exams. Manuscript submitted for publication.
- Meer, N. M., & Chapman, A. (2014). Assessment for confidence: exploring the impact that low-stakes assessment design has on student retention. *The International Journal of Management Education*, 12(2), 186-192.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63(2), 81.
- Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of experimental psychology*, 64(5), 482.
- Murphey, T. (1998). Motivating with near peer role models. *On JALT97: Trends & Transitions*, 201-205.
- Murphy, J., Hofacker, C., & Mizerski, R. (2006). Primacy and recency effects on clicking behavior. *Journal of computer-mediated communication*, 11(2), 522-535.

- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education, 31*(2), 199-218.
- Niehaus, K., & Adelson, J. L. (2013). Self-concept and native language background: A study of measurement invariance and cross-group comparisons in third grade. *Journal of Educational Psychology, 105*(1), 226.
- Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology, 162*, 31-38.
- Ofsted. (2019). *Inspection of Unnamed School*. <https://files.ofsted.gov.uk>
- Ofsted. (2019). *Research for Education Inspection Framework*.
Assets.publishing.service.gov.uk. Retrieved 11 May 2022, from
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/963625/Research_for{EIF_framework_updated_references_22_Feb_2021.pdf.
- Ofsted. (2022). *School inspection handbook*. GOV.UK. Retrieved 11 May 2022, from
<https://www.gov.uk/government/publications/school-inspection-handbook-eif/school-inspection-handbook>.
- Ose Askvik, E., Van der Weel, F. R., & van der Meer, A. L. (2020). The importance of cursive handwriting over typewriting for learning in the classroom: a high-density EEG study of 12-year-old children and young adults. *Frontiers in Psychology, 11*, 1810.

- Pollack, J. M., Najarian, M., Rock, D. A., & Atkins-Burnett, S. (2005). Early Childhood Longitudinal Study, Kindergarten Class of 1998? 99 (ECLS-K). Psychometric Report for the Fifth Grade. NCES 2006? 036. *National Center for Education Statistics*.
- Redmond, B. F. (2010). Self-Efficacy Theory: Do I think that I can succeed in my work. *Work Attitudes and Motivation*.
- Reid, M. A., & Barrington, H. A. (1999). *Training interventions: Promoting learning opportunities*. Institute of Personnel and Development.
- Roediger III, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011). Test-enhanced learning in the classroom: long-term improvements from quizzing. *Journal of Experimental Psychology: Applied*, 17(4), 382.
- Roediger III, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on psychological science*, 1(3), 181-210.
- Rohrer, D., & Pashler, H. (2012). Learning Styles: Where's the Evidence?. *Online Submission*, 46(7), 634-635.
- Romanelli, F., Bird, E., & Ryan, M. (2009). Learning styles: a review of theory, application, and best practices. *American journal of pharmaceutical education*, 73(1).
- Rosenberg, M. (1965). Rosenberg self-esteem scale. *Journal of Religion and Health*.

- Rosenshine, B. (2012) Principles of Instruction: Research-Based Strategies That All Teachers Should Know. *American Educator*, 36(1), 12-39.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, 18(2), 119-144.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in education: principles, policy & practice*, 5(1), 77-84.
- Schacter, D. L., & Buckner, R. L. (1998). On the relations among priming, conscious recollection, and intentional retrieval: Evidence from neuroimaging research. *Neurobiology of learning and memory*, 70(1-2), 284-303.
- Schoen, I. (2012). Effects of method and context of note-taking on memory: handwriting versus typing in lecture and textbook-reading contexts.
- Schroeder, R. W., Twumasi-Ankrah, P., Baade, L. E., & Marshall, P. S. (2012). Reliable digit span: A systematic review and cross-validation study. *Assessment*, 19(1), 21-30.
- Schwieren, J., Barenberg, J., & Dutke, S. (2017). The testing effect in the psychology classroom: A meta-analytic perspective. *Psychology Learning & Teaching*, 16(2), 179-196.
- Shallice, T., & Warrington, E. K. (1977). Auditory-oral short-term memory impairment and conduction aphasia. *Brain and Language*, 4(4), 479-491.
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of educational research*, 46(3), 407-441.

- Sherrington, T. (2020). *Rosenshine's Principles in Action*. John Catt Educational.
- Simzar, R. M., Martinez, M., Rutherford, T., Domina, T., & Conley, A. M. (2015). Raising the stakes: How students' motivation for mathematics associates with high- and low-stakes test achievement. *Learning and individual differences*, 39, 49-63.
- Smoker, T. J., Murphy, C. E., & Rockwell, A. K. (2009). Comparing memory for handwriting versus typing. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 53, No. 22, pp. 1744-1747). Sage CA: Los Angeles, CA: SAGE Publications.
- Sotola, L. K., & Crede, M. (2021). Regarding class quizzes: A meta-analytic synthesis of studies on the relationship between frequent low-stakes testing and class performance. *Educational Psychology Review*, 33(2), 407-426.
- Srivastava, S. K. (2013). To Study the Effect of Academic Achievement on the Level of Self Confidence. *Journal of Psychosocial Research*, 8(1).
- Stabback, P. (2016). What Makes a Quality Curriculum? In-Progress Reflection No. 2 on "Current and Critical Issues in Curriculum and Learning". *UNESCO International Bureau of Education*.
- Steiner, D. D., & Rain, J. S. (1989). Immediate and delayed primacy and recency effects in performance evaluation. *Journal of Applied Psychology*, 74(1), 136.
- Storm, B. C., Bjork, E. L., & Bjork, R. A. (2008). Accelerated relearning after retrieval-induced forgetting: the benefit of being forgotten. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 230.

- Thomas, C. R., & Gadbois, S. A. (2007). Academic self-handicapping: The role of self-concept clarity and students' learning strategies. *British Journal of Educational Psychology*, 77(1), 101-119.
- Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., & Najarian, M. (2009). Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K): Combined User's Manual for the ECLS-K Eighth-Grade and K-8 Full Sample Data Files and Electronic Codebooks. NCES 2009-004. *National Center for Education Statistics*.
- Underwood, B. J. (1957). Interference and forgetting. *Psychological review*, 64(1), 49.
- Vasquez, K. (2009). Learning styles as self-fulfilling prophecies. *Getting culture: Incorporating diversity across the curriculum*, 53-63.
- Wang, W., Subagdja, B., Tan, A. H., & Starzyk, J. A. (2012). Neural modeling of episodic memory: Encoding, retrieval, and forgetting. *IEEE transactions on neural networks and learning systems*, 23(10), 1574-1586.
- Wheeler, M. A., & Roediger III, H. L. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science*, 3(4), 240-246.
- Wiley, R. W., & Rapp, B. (2021). The effects of handwriting experience on literacy learning. *Psychological Science*, 32(7), 1086-1103.

Appendices

Appendix A

Teacher Focus Group Questions

1. How easy or difficult did you find implementing the quick quizzes into the curriculum? Why was this?
2. How did your class respond to the quizzes?
3. Do you think the quizzes contributed to the progress your class made during the unit?
4. Do you think quick quizzes are a good tool for assessment for learning?
5. Were there any children who could not access the quizzes and why?
6. Did you find the time you spent planning the quizzes was worth the benefit it gave to the children's learning?
7. Do you think the quizzes contributed to the children's self confidence in the topic work?
8. Will you continue to use quick quizzes in your planning? If not, why not?

Appendix B

Self-Description Questionnaire (SDQ)

PUT AN "X" IN THE BOX UNDER YOUR ANSWER				
	Not at all true	A little bit true	Mostly true	Very true
1. I have lots of friends	☐	☐	☐	☐
2. I am good at all school subjects	☐	☐	☐	☐
3. I feel angry when I have trouble learning	☐	☐	☐	☐
4. I get good grades in reading	☐	☐	☐	☐
5. I often argue with other kids.	☐	☐	☐	☐
6. Work in math is easy for me	☐	☐	☐	☐
7. I worry about taking tests	☐	☐	☐	☐
8. I enjoy doing work in all school subjects.....	☐	☐	☐	☐
9. It's hard for me to pay attention	☐	☐	☐	☐
10. I like reading	☐	☐	☐	☐

PUT AN "X" IN THE BOX UNDER YOUR ANSWER

	Not at all true	A little bit true	Mostly true	Very true
11. I make friends easily	☐	☐	☐	☐
12. I cannot wait to do math each day.....	☐	☐	☐	☐
13. Work in reading is easy for me	☐	☐	☐	☐
14. I often feel lonely	☐	☐	☐	☐
15. Work in all school subjects is easy for me.....	☐	☐	☐	☐
16. I get good grades in math	☐	☐	☐	☐
17. I get distracted easily	☐	☐	☐	☐
18. I am interested in reading	☐	☐	☐	☐
19. I get along with kids easily ...	☐	☐	☐	☐
20. I feel sad a lot of the time.....	☐	☐	☐	☐

PUT AN "X" IN THE BOX UNDER YOUR ANSWER

Not at all true	A little bit true	Mostly true	Very true
-----------------	-------------------	-------------	-----------

- | | | | | |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| 21. I cannot wait to read each day | ☐ | ☐ | ☐ | ☐ |
| 22. I am interested in math | ☐ | ☐ | ☐ | ☐ |
| 23. It's hard for me to finish my school work | ☐ | ☐ | ☐ | ☐ |
| 24. I am easy to like | ☐ | ☐ | ☐ | ☐ |
| 25. I worry about doing well in school | ☐ | ☐ | ☐ | ☐ |
| 26. I can do very difficult problems in math..... | ☐ | ☐ | ☐ | ☐ |
| 27. Other kids want me to be their friend | ☐ | ☐ | ☐ | ☐ |
| 28. I like all school subjects | ☐ | ☐ | ☐ | ☐ |
| 29. I worry about finishing my work..... | ☐ | ☐ | ☐ | ☐ |
| 30. I like math | ☐ | ☐ | ☐ | ☐ |

	Not at all true	A little bit true	Mostly true	Very true
31. I have more friends than most other kids.....	☐	☐	☐	☐
32. I worry about having someone to play with at school.....	☐	☐	☐	☐
33. I am good at reading	☐	☐	☐	☐
34. I get in trouble for talking and disturbing others	☐	☐	☐	☐
35. I like reading long chapter books.....	☐	☐	☐	☐
36. I enjoy doing work in math ...	☐	☐	☐	☐
37. I get in trouble for fighting with other kids	☐	☐	☐	☐
38. I look forward to all school subjects	☐	☐	☐	☐
39. I enjoy doing work in reading	☐	☐	☐	☐
40. I feel ashamed when I make mistakes at school.....	☐	☐	☐	☐

Appendix C

Pupil Focus Group Questions

1. Did you enjoy the quick quizzes at the start of your topic lessons?
2. Do you think the quick quizzes helped you to remember the information you learnt in the topic?
3. Do you think the quick quizzes helped your self confidence in the topic?
4. Would you have changed anything about the quizzes? If so, what?
5. Would you like the quick quizzes to continue?

Appendix D

Headteacher Letter of Consent

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

Headteacher

----- Primary School

6th November 2021

Dear ----,

I am writing to inquire about conducting research in school this academic year. As you know, I am studying for the Master's in Learning and Teaching at Oxford University, supervised by Dr ----- . In my final research project "The effectiveness of recall strategies in the primary classroom to increase pupils' self-confidence in knowing and remembering more" I will explore how we can embed current recall techniques in a primary classroom and measure their effectiveness in allowing the pupils to recall more of their knowledge.

The research will take place across KS2. At the start of Spring 1 half term, all pupils would need to complete a short questionnaire about their confidence at school. There are 42 agree or disagree questions and should take approximately 10 minutes of class time. A copy of the questionnaire can be found here.

Then, over the half term of foundation subject work, one class per year group would complete a low-stakes quick quiz at the start of the lesson in a written format, whereas the other class would complete it using discussion. At the end of the unit of work, the children will be assessed in the usual way. They will complete the confidence questionnaire again, and their scores compared to their teacher assessments. I would also host pupil and teacher focus groups to gather opinions on the quick quizzes and how they have impacted pupil progress and confidence.

By participating in the research, the school would be contributing to a project that will deepen our understanding of recall strategies in the primary classroom and how these can be used to increase academic pupil self-confidence. I hope to conduct this research between January 2022 and June 2022, with the main bulk of data collected in Spring 1 half term, but with flexibility for extra data to be produced until June if it is needed.

Oxford University has strict ethical procedures on conducting 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, teachers 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, through: research.office@education.ox.ac.uk.

I look forward to hearing from you.

Yours sincerely,

The effectiveness of recall strategies in the primary classroom to increase pupils' self-confidence in knowing and remembering more.

- 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.

Headteacher's signature

Please return this form to me.

Thank you for your help.

Appendix E
SDQ Scores

Reading	Reading	Maths	Maths	Maths	Academic	Academic	Academic	Social	Social	Social
End	Difference	Start	End	Difference	Start	End	Difference	Start	End	Difference
2.61	-0.43	2.93	2.52	-0.41	2.73	2.29	-0.44	2.69	2.22	-0.47
2.51	0.04	2.99	2.91	-0.08	2.51	2.22	-0.29	2.46	2.35	-0.11
3.1	0.04	3.23	3.18	-0.05	2.65	2.58	-0.07	2.39	2.31	-0.08
2.57	-0.12	2.63	2.59	-0.04	2.29	2.08	-0.21	2.35	2.21	-0.14
2.65	-0.3	2.72	2.42	-0.3	2.37	2.09	-0.28	2.26	2.14	-0.12
2.64	0.24	2.91	3.07	0.16	2.38	2.42	0.04	2.42	2.2	-0.22
2.37	-0.38	2.79	2.51	-0.28	2.11	1.96	-0.15	2.26	2.24	-0.02
2.6	-0.21	3.11	3.05	-0.06	2.63	2.44	-0.19	2.68	2.51	-0.17

Class	Total	Total End	Total	Reading
	Start		Difference	Start
Year 3	125.33	106.59	-18.74	3.04
Written				
Year 3	112.48	105.59	-6.89	2.47
Oral				
Year 4	120.86	115.72	-5.14	3.06
Written				
Year 4	109.41	105.78	-3.63	2.69
Oral				
Year 5	112.12	101.35	-10.77	2.95
Written				
Year 5	108.21	107.32	-0.89	2.4
Oral				
Year 6	108.07	99.11	-8.96	2.75
Written				
Year 6	118	108.64	-9.36	2.81
Oral				

Note. Positive score increases have been highlighted.