Revising revision: Encouraging self-testing behaviours when preparing for high-stakes mathematics examinations

Joanna Hall

A Research & Development Project Submitted for the MSc in Learning & Teaching 2019

DEPOSIT AND CONSULTATION OF THESIS

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

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

Surname	Hall
First Name	Joanna
Faculty Board	Education
Title of Dissertation	Revising revision: Encouraging self-testing behaviours when preparing for high-stakes mathematics examinations

Declaration by the candidate as author of the dissertation

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

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

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

REVISING REVISION:

Encouraging self-testing behaviours when preparing for high-stakes mathematics examinations

Contents

Chapter one – Introduction

1.1 – Introduction……7

1.2 – Rationale and context……8

1.3 – Purpose and aims……10

Chapter two – Literature review

2.1 – Introduction……11

2.2 – The rise of a knowledge based curriculum……11

2.3 – The influence of cognitive science on education……13

2.4 – The testing effect……16

2.5 – The transfer of knowledge……21

2.6 – Effective and ineffective study techniques……25

2.7 – Self-regulation, grit and creating desirable difficulties……31

2.8 – Harnessing change……38

2.9 – Conclusions from the literature……40

2.10 – Research questions……41

2.11 – Hypothesis……39

Chapter three – Methodology

3.1 – Introduction……42

3.2 – Methodology……42

3.4 – Intervention……44

3.4.1 – Outline of low-stakes tests……44

3.4.2 – Participants of study……45

3.4.3 – Outline of collaboration……46

3.4.4 – Introduction to the project……47

3.5 – Data sources……47

3.5.1 – Questionnaires……44

3.5.2 – Semi-structured interviews – Students……49

3.5.3 – Semi-structured interviews – Teachers……51

3.5.4 – Mock examination data……51

3.6 – Ethical considerations……52

3.6.1 – Confidentiality……53

3.6.2 – Consent……53

3.6.3 – Power imbalance……54

3.6.4 – Protecting data……55

3.6.5 – Avoiding disadvantage……55

3.6.6 – Low-stakes tests……56

Chapter four – Findings and discussion

4.1 – Introduction……57

4.2 – Curriculum content……57

4.3 – The testing effect……60

4.4 – Transfer of knowledge……62

4.5 – Initial study techniques……64

4.6 – Self-regulation, grit and creating desirable difficulties……69

4.7 – Harnessing change……73

Chapter five - Conclusions

5.1 – Conclusions……...……...…..79

5.2 – Limitations of the project……...……82

5.3.1 – Time……82

5.3.2 – Questionnaires……..…..……82

5.3.3 – Sample size and population……..……..82

5.3 – Implications for future practice……...…...……84

5.4 – Closing remarks……..……...…….85

List of appendices

Appendix 1: Example of low-stakes test

Appendix 2: Pre-intervention student questionnaire

Appendix 3: Post-intervention student questionnaire

Appendix 4: Lower ability pre-intervention study strategy questionnaire

Appendix 5: Higher ability pre-intervention study strategy questionnaire

Appendix 6: Lower ability post-intervention study strategy questionnaire

Appendix 7: Higher ability post-intervention study strategy questionnaire

Appendix 8: Powerpoint for introducing the intervention to students

Appendix 9: Grit questionnaire

Appendix 10: Questions for student interview pre-intervention

Appendix 11: Questions for student interview post intervention

Appendix 12: Questions for semi-structured teacher interview

Appendix 13: CUREC ethical approval form

Appendix 14: Letter of approval to head teacher

Chapter One – Introduction

1.1 – Introduction

Giving students the necessary tools to become lifelong learners can be a challenging task for any teacher. With the 2017 introduction of the newly reformed GCSE curriculum, secondary school students have been faced with tougher examinations encompassing a higher standard of rigour and an increase in curriculum content. Mathematics in particular has historically been found to be one of the most intellectually demanding courses with regards to the volume of what needs to be learnt in order to be successful.

With increased demands on teachers to deliver more content in the same amount of time, problems now arise. Although Education Secretary at the time, Damien Hinds, claimed in April 2019 that the reform to the new GCSE exams reduced anxiety and stress (Sunday Times, 2019) by removing the majority of coursework in place of exams, a poll carried out by the National Education Union painted a contrasting picture. With 650 secondary schools taking part in their 2018 survey, the NEU claimed that 89% of students believed to suffer from increased levels of anxiety and stress as a result of the higher demands of the new GCSE's and a larger emphasis to perform in exams, a story replicated not just nationally, but within the confinement of my own school too. As a result of the overhaul, students' revision has been put under increased strain, and the importance of its effectiveness is now critical in order to accommodate drastic increases of content across all subject areas.

Although preparation for high stakes examinations has always been of importance to key stage 4 students and teachers alike, the new Ofsted framework (2019) has placed

additional weight on the significance of creating learning environments where long-term learning is at its heart. The report suggests it is fundamental to student learning that educators integrate new knowledge regarding techniques to improve long-term memory, and focus on making lasting learning connections that promote fluid understanding.

When executed well, effective self-study time can provide a substantial improvement to the final grade or mark that a student receives (Karpicke, Butler and Roediger, 2009). On the opposing end of the scale, unproductive revision time could equate to hours of wasted, valuable time in the run-up to examinations. What is concerning is the high proportion of students and teachers who are seemingly unaware of how effective or ineffective some study strategies might be, with some even undermining student achievement (Dunlosky et al, 2013). According to the BBC (Cohen, 2013), teachers are lacking in awareness of how memory really works, and as a result are not advising students on the most effective ways to use their valuable revision time. It is therefore the role of the practitioners who surround them to encourage effective self-study behaviours in unsupervised environments. What has become problematic is the juxtaposition between the disproportional amount of teacher guidance when it comes to curriculum knowledge, and a lack of support schools and teachers offer where self-study and revision techniques are concerned.

1.2 – Rationale and context

Since 2016, a downward annual trend of GCSE examination results in mathematics has led to a bleak prediction for the 2019 results. As a newly appointed head of mathematics and an Ofsted inspection imminent, logical and well-informed

measures, driven by research and the individual needs of the students are required in order to drive up attainment.

Up until this date, advising students of how best to revise has been delivered sporadically at my school. Little guidance is given until key stage 4, with the biggest emphasis placed on year 11. Although my school has taken steps to address informing students of effective revision techniques, one-off seminars in year 11 and sparse, in-lesson advice weeks before high stakes examinations can have little to no effect self-study choice, especially if bad habits have already set in. This issue was highlighted when an external company claiming to boost memory by using techniques such as story telling and mind mapping, visited the school to deliver a one-off whole day intervention to our year 10 students. In hindsight, the enthusiasm witnessed was more likely due to a day off timetable, rather than being eager to try out their newly acquired memory tricks to prepare for upcoming tests. This was a costly strategy for the school, which resulted in a post intervention questionnaire 6 months on, carried out by the senior leadership team, which revealed 93% of the 258 students who attended the workshop did not use any of the strategies that were delivered on the day, as they prepared to sit their end of year mock examinations. But perhaps an even more worrying statistic uncovered that 87% of the year 10 students at the time confessed to not knowing how to revise properly for their exams.

This trend has also been noted in younger year groups. Students lower down the school openly admit that they do not know how to revise material that has been delivered in lesson time, and will often choose not to study for topic or end of year tests. This has therefore paved the way for bad habits to be developed by the time critical stages schooling are reached, which further adds to the tension and stresses

students experience as they approach their GCSE examinations, creating pockets of disengagement. The rationale for my project has subsequently been influenced by an increasing number of students who are seemingly struggling with self-study, in times of increased pressure as a result of tougher examinations where effective revision can contribute to higher success rates.

1.3– Purpose and aims

The principle aims of my research was to test whether a demonstration of robust revision strategies in lesson time has the ability to not only drive up progress in high-stakes examinations, but to educate students of the many benefits self-testing strategies can have and the potentially damaging effects less fruitful strategies entail.

With some studies suggesting that lower achieving students are more likely to adopt less robust study strategies than their higher achieving peers (Gurung, 2005), and others indicating all students are at a disadvantage because teachers revision strategy advice is ill-informed (Dunlosky et al, 2013), it comes as no surprise that some students are not embracing effective techniques that have been proven to increase long-term retention of knowledge. In light of this, I will compare the differences in initial study choices between higher and lower ability students at my school, and analyse these against any changes that occur because of the intervention. I also intend to discover the views and opinions of students in year 11 and teachers of that year group regarding self-study strategies, and how best to support students not just in the lead up to high-stakes examinations, but throughout their secondary school lives.

Chapter Two – Literature Review

2.1 - Introduction

To begin, I examined the influential factors that played a part in the reform of the GCSE examinations in 2017. The focal point of testing to enhance recall of knowledge in high stakes examinations, and how this may impact on higher and lower ability students, is then paired with the idea of how students choose to allocate their self-study time. Of particular interest to this research will be levels of engagement with methods such as retrieval practice in addition to motivational factors that influence strategy choice of both higher and lower ability students.

2.2 - The rise of knowledge based curriculum

In 2009, the Minister of Education at the time, Michael Gove, set out desires to overhaul the secondary school curriculum in England in favour of one that focussed fundamentally on a core body of knowledge. Gove was quoted at a speech to the Royal Arts Society in 2009 that the overhaul plans in place were 'to ensure that the acquisition of knowledge within rigorous subject disciplines is properly valued and cherished' and that by doing so would not only benefit the education system, but would be advantageous for social mobility. He concluded that the English education establishment had taken a flawed direction, and its foundations were set in 'romantic ideologies' with approaches such as personalised learning traits having gained much momentum, despite limited conclusive evidence.

Gove's intentions were influenced by his readings of American educator and academic literary critic E. D. Hirsch. Hirsch's (2006) views on the lack of success the US education system had experienced hit a notion that is now widely

acknowledged, citing one of the fundamental weaknesses of a substandard education system is where personalised learning, tailored to the specific needs of a child, takes precedent over the developing a core understanding of 'cultural literacy' as a principle value. Christodolou, (2013) takes this idea further, and refers to the personalisation of the curriculum as a myth that has no sound evidence to support its claim that better understanding is achieved. She maintained that educators had a fixated vision on personalising curriculum content to meet the needs of individual learners (Christodolou, 2013). In 2014, a survey revealed that 93% of teachers in the UK agreed with the statement 'Individuals learn better when they receive information in their preferred learning style (for example, visual, auditory or kinaesthetic)' (Howard-Jones, 2014, p.2), despite no conclusive evidence to support such claim.

2010 saw the Conservative party form a coalition government, and true to his word, Gove set out plans to overhaul the curriculum, and released an initial first draft of the reform in 2013. What followed was a thorough and rigorous overhaul of the English secondary school examination system, whereby Mathematics examinations in 2017 contained considerably more content than the previous year, and as such, placed knowledge at the core of the curriculums intentions. Demands on recall and domain specific knowledge were put under additional pressure as a result of an increase in examinations to assess students and a decrease in coursework. Initial concerns were raised over lower achieving students (Trkulja, 2013), arguing that the exam system would do nothing but further segregate those who were academically less advanced - precisely the opposite view of Hirsch's (1987) model of social fluidity.

With large changes imminent, educators had to tread carefully around evidence being fed to them and what they believed to be correct. Between the years of 1967 and 1995, the largest and most costly education study in history, entailing over 70,000 students across America compared a variety of teaching approaches that included direct instruction and a cognitively orientated curriculum (Stallings and Kaskowitz, 1974). In all areas, direct instruction proved to be the most effective in improving both mathematics and literacy skills. So why did it take until 2017 for changes to the curriculum to be put in place to reflect the advantages of what Hirsch (2006) describes as common cultural literacy? And what are the implications on the way students learn as a result of an increased examination assessment model and a decrease in coursework? Would an increase in rigour and content call for better understanding of how our minds really work, rather than relying on dated and inaccurate teaching techniques?

2.3 - The influence of cognitive science on education

With heavily increased demands on curriculum content, it becomes imperative that students make best use not only of contact time whist at school, but also implement effective strategies in unsupervised environments in periods of self-study. Cognitive science may hold significant implications in the realms of education (Heritage, 2018), and practitioners would be wise to use research in this domain, rather than outdated myths that allude to personalisation (Christodolou, 2013) to shape the advice and strategy demonstrations encompassing self-study techniques that are ultimately cascaded down to students.

The roles that both learning and memory play are inherently linked to long-term retention, and a highly intricate structure, beyond the level of what secondary school educators are expected to be at ease with, still remains unclear to experts in this field (Baddeley, Eysenck and Anderson, 2009). One of the central factors contributing to an enhancement in mathematical skill is the proficiency of working memory. Miller et al (1960), defines working memory as the cognitive arrangement of ideas and experiences that gives people the ability to make such information readily available to use in new contexts. Engle and Kane (2004) take this notion further, and cite the benefits of storing large amounts of background knowledge in long-term memory, is that said knowledge is not taking up valuable capacity in our fragile working memories. Furthermore, (Baddeley, 2012) argues that more complex cognitive activities are made easier if we have access to a higher capacity of working memory.

Over the last 20 years, extensive research into factors contributing to a range of ability level has cited the capacity of working memory as a major benefitting factor (e.g. Bull and Scerif, 2001; Geary, Hoard, Byrd-Craven and DeSoto, 2004). This, in turn, has led researchers to use working memory capacity and regulation as a predictor of future success in mathematics (LeFevre et al, 2005), and forecast that students who possess larger working memory capacity are in a more educationally desirable position than those who do not (Swanson and Jerman, 2006).

Notably, variances between students with lower and higher working memory capacities and their ability in mathematics have not been found to be consistent (Censabella and Noël, 2008), making it difficult for educators to cascade educational research into tangible teaching practice. As such, a one size fits all strategy is

difficult to implement to meet the needs of all learners across the ability spectrum. It has therefore been speculated that lack of mathematical ability is subject specific (Bull and Scerif, 2001). These important findings further support Hirsch's (2006) assertion that a broad body of core knowledge is how best to tackle the complications that arise from educating students with different levels of working memory capacity, rather than focussing on the 'romantic ideologies' that Gove (2009) speaks of.

Earlier research (McGeoch, 1932) argued that knowledge and memories would always remain stored in our long-term memories, but they can become inaccessible. McGeoch (1932) alludes to the notion that information committed to long-term memory can be replaced by new and competing information, making previous knowledge inaccessible. In contrast, Bjork and Bjork's (1992) theory of disuse hypothesises that forgetting is advantageous to the learning process. They challenged the perception that forgetting is damaging to the learning process, explaining that items held in memory each have a storage strength, (how easy information is to remember in the present), and a retrieval strength, (how easy information is to remember in the future). Principally, Bjork and Bjork (1992) state that human memory has a limitless capacity, and go on to explain is retrieval strength is improved as the repeated revisiting of that information over regular, spaced intervals. Fundamentally, students are to engage in numerous exposures of material for it to be converted from working memory to long-term memory (Willingham, 2009).

The notion of repeated remembering and forgetting is by no means a new phenomenon, and as far back as 1885, Ebbinghaus developed his forgetting curve

through inferences made from recalling learnt information. Although doubts have been raised over the validity of this specific experiment (Roediger, 1985), as his theory was conducted on a sample size of one – himself, spacing learning is now a widely accepted and replicated procedure in the field of cognitive science as a means of converting knowledge from working memory into long-term memory (Joeri and Jaap, 2015).

In turning to cognitive science to better understand our minds, it must be noted that advice within these realms must be met with caution. The vast majority of studies into improving cognitive capacity have been carried out in controlled laboratories (Skinner, 1985; Willingham, 2009), a variable that we simply cannot replicate within the realms of this research project. Care must therefore be taken when disseminating research from within the fields of cognitive science. That being said, research within the cognitive science domain serves as an additional layer of understanding to practitioners, and large educational gains may be seen if this advice is used in the correct way in school settings (Dubinsky, Roehrig and Varma, 2013).

2.4 - The testing effect

For Hirsch (2006) the use of tests are twofold. The most advantageous method of securing the required knowledge he speaks of, whilst holding schools accountable for the progress of students under their care is to regularly assess through testing, in order to uphold the 'twin aims of equality and fairness' (p.91). But in recent years, a surge in evidence has made claims of the many other benefits tests can have, of which most are surprisingly unanticipated by both students and teachers, given that tests are common practice in most classrooms. With students facing increased levels

of curriculum content, and a heavier emphasis on examinations at the end of their GCSE course, could tests serve as a learning tool, rather than purely to assess understanding?

There is an increasing body of evidence that suggests that testing has an extensive range of additional advantages that are overshadowed by the perception that the principle purpose of testing is to assess intelligence. If students and teachers are able to detach from the assessment habits that can govern many classrooms (Black and Wiliam, 1998), tests as a instrument to enhance learning could serve as a powerful tool. Karpicke and Roediger (2008) are in agreement with this view, and report that the current perspective of testing students is a 'neutral event that merely measures learning' (p.966). Roediger et al (2011) add to this argument through a collection of compelling evidence. They set out to describe the many other benefits testing can have as an aid to improving long-term memory and retention of knowledge, whilst being viewed by both students and teachers as an undervalued and underused strategy to assist in the long-term acquisition of knowledge.

Testing has been found to increase metacognitive monitoring (Shaughnessy and Zechmeister, 1992; Kornell and Son, 2009) by identifying gaps in knowledge (Amlund, Kardash, and Kulhavy, 1986), can lead students to learn more from subsequent learning episodes (e.g. Izawa, 1966; Karpicke and Roediger, 2007; Pyc and Rawson, 2010), and produce a higher quality of organisation of knowledge (e.g. Gates, 1917; Masson and McDaniel, 1981; Congleton and Rajaram, 2011).

However, arguably the most crucial advantage at secondary school education level are reported in studies such as Wheeler and Roediger (1992) and Wheeler, Ewers and Buonanno (2003) who revealed the retrieval of information through testing can lead to better performance in later tests by slowing the rate of forgetting, forcing participants to use memory retrieval to assist in reinforcing the retrieval strength of a memory (Bjork and Bjork, 1999). Specifically in Wheeler and Roediger (1992), reports that studying material once and sitting tests on the studied material three times will produce an 80% improvement in retention of the studied material than if participants were to study three times and test just once.

Importantly, the Dunlosky et al (2013) review of ten commonly used self-study strategies, many of which will be discussed later in this chapter, rated self-testing as one of the most highly effective strategies. Dunlocky et al (2013) stated that the testing effect as an approach to revision had 'been demonstrated across an impressive range of practice-test formats, kinds of material, learner ages, outcome measures, and retention intervals' (p.35). In addition to the diagnostic benefits that testing can pose on a formative level (Kornell and Bjork, 2007), and the advantage that exposing students to effective learning strategies in lesson time will promote their use as a self-study strategy (EEF, 2018), what results is a powerful teaching tool (Didau, 2013).

Differing opinions of testing as an aid to enhance memory may argue that learning time is being sacrificed, and therefore deep understanding is at the expense of retrieval practice. As teachers, we are divided between teaching students knowledge required for deep understanding and connection of ideas presented in the curriculum,

and being able to pass examinations. Skemp (1976) appreciated that learning time is split between teaching students what to lean and how they are to learn it and as such, a fine equilibrium must be met. Skemp (1976) takes this idea further, and argues the case that there are two types of understanding: relational and instrumental. His definition of relational understanding not only underpins what students do, but why they are doing it. Instrumental understanding is what Skemp (1976) describes as 'rules without reasons' (p.2), and alludes to notions of rote learning without sound understanding of the content. Skemp (1976) is of the opinion that relational understanding is what teachers should aspire for their students, with instrumental understanding presenting little in that way of making connections and inferences between mathematical topics.

That being said, this can present difficulty for teachers. With intense pressure on teachers to deliver results, coupled with time constraints following the increased demand in rigour and content to the new style GCSE's from 2017, it could be argued that techniques such as retrieval practice could narrow the relational understanding of students by not allowing them to explore connections between topics.

Further criticisms of the literature that demonstrate the testing effect surfaced more recently, and disputed that testing students on solitary fluency skills, rather than tackling complex problem solving questions similar to those seen in an exam, encourages students to solely commit facts to memory, rather than reason between different aspects of a subject (Van Gog and Sweller, 2015). Furthermore, Van Gog and Sweller (2015) claimed the testing effect was absent when participants attempted to retrieve complex material.

But these claims were abruptly quashed by a counter argument from Karpicke and Aue (2015) declaring that the literature reviewed in Van Gog and Sweller's (2015) study had failed to acknowledge several investigations that indicated a positive testing effect when retrieving complex material. A further look back into the literature from McDaniel et al (2013) also disputed this view, claiming that testing was able to encourage deeper learning beyond the surface of a specific answer. Could it therefore be said that the testing effect could serve as a tool for striking a balance between assisting in lasting learning, whilst avoiding taking too much valuable curriculum time?

With literature surrounding retrieval practice suggesting large benefits in an array of contexts and educational settings, attention should now turn to the limitations of the testing effect. Recent research by Agarwal et al (2017) into sub groups of students who benefitted more from delayed testing to assist working memory discovered that those who were classed as having lower working memories gained a significantly greater benefit than those who were classed as having higher working memories. Agarwal et al (2017) concludes if this is the case, their findings deliver pivotal educational repercussions for improving the learning circumstances for lower ability students, serving these groups of learners a tool to 'level the playing field' (p.770) by placing them in an educationally more desirable position (Swanson and Jermam, 2006).

Given evidence presented by Agarwal et al (2017), surely it would be the case that lower ability students were adopting this strategy in order to ease pressure on working memory and reap the educational rewards of such a phenomenon? The

literature suggests this may not be the case. It has been reported that students who are least likely to adopt testing as a self-study strategy are those who are less metacognotivley aware of the benefits testing can have (Einstein, Mullet and Harrison, 2012),. Moreover, those who do take up testing as a regular self-study habit, reportedly do so to exploit the diagnostic benefits, rather than consciously utilising self-testing as a method to improve long-term memory (Einstein, Mullet and Harrison, 2012), Seemingly, students are unaware of the vast favourable consequences that regular retrieval practice possesses.

2.5 – The transfer of knowledge

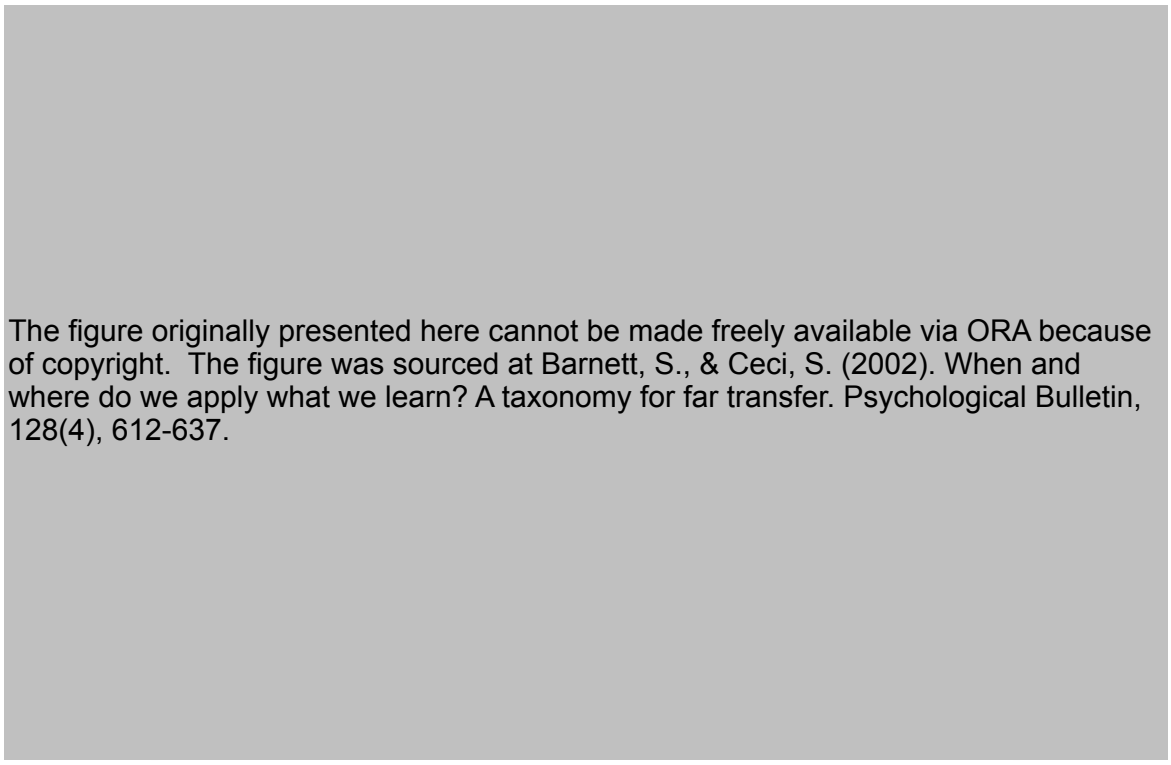
What we have discussed so far, is the forgetting of information can be advantageous to the learning process (Bjork and Bjork, 1992), and regular, spaced testing can assist in the process of repeated forgetting and memorising of material so it can be committed to our long-term memories in order to free up working-memory space (Engle and Kane, 2004). Proof of the testing effect has been present in a variety of formats, across a range of material of varying complexities (McDaniel, 2013) to paint an impressively broad and promising picture for secondary school practitioners and students (Dunlosky et al, 2013). But we must look at why we would implement such a strategy into our classrooms. Of course, our priority is to build lifelong learners, where students remember what we teach them, weeks, months and even years after they have left us, rather than only in the short term, but we must not be ignorant of the fact that the success of a student beyond secondary education lies largely at the hands of good examination results. For the testing effect to have any impact on how students perform in their high-stakes examinations, they must be able to transfer said knowledge into the exam hall with them. However, some might

question if we are doing students a disservice by teaching them to pass exams, rather than to gain a deep understanding of the subject (Skemp, 1976).

For Willingham (2002), knowledge can only be classed as flexible when it has the ability to be administered to new and unfamiliar situations, separate to the condition in which the knowledge was acquired. This alludes to the thought of the standard of the flexibility, rather than the standard of the knowledge itself in the realms of a fixed state. With more recent studies from Roediger et al (2011) concurring, adding that testing assists in inducing 'readily accessible information that can be flexibly used to solve new problems' (p.14).

Early research carried out by Thorndike and Woodworth from (1901) gave rise to his identical elements theory of transfer, which stipulates that the degree to which knowledge is acquired from one circumstance to another is limited by the likeness between the two circumstances. The more familiar a situation is, the more potential there is to transfer knowledge and information. In the same respect, if two situations have no similar features, knowledge learned in one situation will not hold any value to the other (Thorndike and Woodworth, 1901). This was demonstrated by a study by Anderson, Reder and Simon (1996), who came to learn that transfer to new situations is dependent on the context the knowledge was initially learnt in. Their study revealed that children living on the streets of Brazil were able to carry out complicated mental mathematical sums when securing sales, but when removed from their street sale environment and asked to perform the same skills presented in a typical school context, failed to demonstrate the same level of expertise.

When casting these notions over an educational setting, the same outcomes are observed. Wooldridge et al (2014) saw students successfully being able to answer questions that were closely related to previously learned subject matter, but found it difficult to use their acquired knowledge in different topic of the same subject. As such, in order to make sense of knowledge transfer, and to clarify meaning for the purpose of this discussion, Barnett and Ceci (2002) formulated their taxonomy of transfer, which attempts to draw clarity by describing the scope to which the transfer of new knowledge or skills could be measured against. According to their model, near transfer refers to the transfer of knowledge and skills to novel situations that share some underlying similarity to the learning context (Barnett and Ceci, 2002; Meijer and Riemersma, 2002). Their taxonomy of transfer sheds light on to a complex web of dimensions that are discussed, including areas such as knowledge domain and temporal context. Far knowledge transfer in such domains supports what Skemp (1976) describes as relational understanding, and applying the taxonomy to previously discussed literature in the context of the testing effect, will provide the discussion with a more rounded awareness of the factors that may influence knowledge transfer.



When Barnett and Ceci's (2002) taxonomy is administered to some of the research that makes bold claims about the rigour of the testing effect, it presents new awareness. Karpicke and Roediger's (2008) study reveals evidence in the temporal and modality contexts of a near transfer of knowledge. The investigation of Agarwal et al (2017) demonstrates limited far transfer characteristics, with near transfer observed on the knowledge domain, physical contexts and modality. Agarwal et al's (2017) study compared retention of knowledge over a 10-minute and two-day period, both of which, according to Barnett and Ceci (2002) are classed as near in transfer. These examples of the testing effect, when Barnett and Ceci's (2002) taxonomy of transfer is applied to them, present limits to whether the testing effect can be used as a method to increasing the transfer of knowledge to unfamiliar environments. As such, this may raise questions as to whether content testing in the classroom has the ability to be transferred to the exam hall.

Despite these observations, a closer look at the literature on the testing effect with regards to the successful far transfer of learning, and, perhaps a pivotal development regarding the far transfer of knowledge, is the finding that retrieval practice as a learning strategy has the power to increase the transfer of knowledge to unfamiliar situations (Agarwal, Bain and Chamberlain, 2012; Butler 2010). In Butler's (2010) experiment, students were given the task of studying written passages of text, on subject matter such as bats. The participants then sat tests on some of the passages, and restudied other parts of the test. A week after the initial learning session, the participants were asked to answer questions in unfamiliar subject domains, and were deemed to be more successful in transferring knowledge to unfamiliar domains if the knowledge had been tested rather than restudied.

Although this cannot be regarded as conclusive, it would be welcome research regarding the practicalities of incorporating retrieval practice in a secondary educational setting. If testing can assist in both the near and far transfer of learning to new contexts, this would surely give added leverage to advising teachers and students to adopt testing as a retention strategy both in and outside the classroom.

2.6 - Effective and ineffective study techniques

Having explored what constitutes effective and long-lasting learning (Karpicke and Roediger, 2008) and desirable difficulties (Bjork, 1994) by taking influence from research from metacognitive awareness (Zimmerman, 1995), it would now be useful to examine what self-study strategies students may be adopting, and whether effective strategy use holds any correlation between high and low achievers.

Furthermore, is there reason behind student decisions not to adopt effective self-study strategies?

In 2013, Dunlosky et al conducted a meta-study in order to determine what specifically constitutes an effective self-study strategy. Dunlosky et al (2013) reported on ten commonly used self-study techniques used by students, and reported on their utility. A high-utility strategy was deemed to be one that could be utilised across a range of abilities and contexts and promoted lasting learning. For the purposes of this research, I will discuss the strategies that are used often in mathematics as a self-study strategy, and will help me answer my research questions: self-testing, use of flashcards, highlighting and underlining, and rereading.

We have previously discussed the benefits that self-testing can have in an educational setting (Karpicke and Roediger, 2008) and with much evidence to support claims that retrieval practice can support lasting learning by assisting long-term memory (Bjork, 1994) it is little surprise that Dunlosky et al (2013) rate self-testing as a revision strategy that carries high utility.

Further to self-testing, flashcard use was rated by Dunlosky et al (2013) as high-utility, as when used correctly, they harness the testing effect whilst distributing practice over time. That being said, it is often the case that students who report using flashcards, use them to re-read information presented on them, rather than to take full advantage of their testing effect potential (Hartwig and Dunlosky, 2012).

The self-study strategies that Dunlosky et al (2013) reported on as having low utility include highlighting, underlining and re-reading, but these strategies are seemingly

popular among students (e.g. Bell and Limber, 2010; Lonka, Lindblom-Ylänne, and Maury, 1994; Nist and Kirby, 1989). And yet despite its popularity as a study strategy, highlighting has been shown to have little effect with regards to long-terms retention of knowledge across an array of different contexts, including children (Rickards and Denner, 1979) and undergraduates, (Todd and Kessler, 1971) and even air force pilots (Stordahl and Christensen, 1956). Oakley (2015) claimed the physical action of highlighting creates an illusion of knowing, luring students into a sense of familiarity with material by thinking they have mastered the text they have highlighted, when in fact they have not.

Rereading, a strategy linked to low levels of learning (Dunlosky et al, 2013) is another strategy that is seemingly popular among students. Rinehart (1986) reported a total of 84% of students revealed they studied for college tests by rereading. Bjork and Bjork (2011) go as far as to say that nothing 'could be further from the way we actually learn and remember' (p.62). Rereading, highlighting and underlining are comparatively straightforward processes and require little in the way of effort that desirably difficult processes possess (Bjork, 1994). Familiarity of material gives the illusion to students that they are engaging in self-study, when the reality is that very little in the way of long-term learning is taking place (Callender and McDaniel, 2009). Instead, students become engaged in the type of mindless, rote practice that Skemp (1976) speaks of, without attempts to construct connections, and when students are asked to grapple with understanding and meaning of text, rereading and underlining does little to produce lasting learning (King, 1992; Mayer, 1996).

Worryingly, Karpicke and Roediger (2008) revealed that when asked to assess the successfulness of various learning strategies, students were unaware of the benefits testing can have, and for this reason, would often favour less robust strategies over techniques that incorporated elements of retrieval practice. McCabe (2011) are in agreement with this view, discovering that students regularly lack understanding that engaging with self-testing behaviours are likely to strengthen long-term learning. With some studies suggesting specific techniques affect performance in high-stakes examination (Bol, Warkentin, Nunnery, and O'Connell, 1999), whilst others advising a wide range of strategy use is optimal (Rovers et al, 2018; Hadwin and Winne, 1998), not only can it be challenging for students to optimise their learning in self-study periods, but it can be difficult for educators to disseminate sound advice grounded in robust research.

A widely replicated opinion shared between both researchers (e.g. Annis and Davis, 1978), and educators alike, is that students do not always employ effective self-study strategies in unsupervised environments. Dunlosky et al (2013) cite one of the reasons for this is that students were reported to gravitate towards self-study strategies that are easier to undertake, are less time consuming, less complex and demand little effort from the student. With Bjork (1994) reminding us that more effortful retrievals are beneficial to creating lasting memories, and the finding that many learners report an improvised approach to their revision technique choice (Kornell and Bjork, 2007), it is little surprise that some students often opt for easier strategies if they are unaware of the benefits or potential damage some techniques may carry.

In addition to the lack of learning that takes place when using strategies such as underlining, highlighting and rereading, caution must be taken over the undercurrents of illusions of competency, a phenomenon first reported by Glenberg, Wilkinson and Epstein (1982). Demonstrated clearly by Newton's (1990) study, participants were asked to predict the probability that an audience will correctly recognise a well-known song that they were to tap out. The individuals given the task of tapping predicted an average of a 50% success rate. The actual success rate was under 3%. This experiment pinpoints the struggle people and students may face given a position of privilege and familiarity, resulting in an overestimate of their actual capabilities. As a result, illusions of competence tend to occur when student's judgment of their understanding of content is biased by information available during study but not available during testing (Jacoby et al, 1994). A student who has read how to solve a quadratic equation may think they know how to carry out the procedure given their perspective bias and immediate familiarity with the content, but in reality, when the time comes to performing the procedure in a test, they fall short.

Of particular importance to this research are the types of students who are more inclined to adopt testing behaviours in a period of self-study. We have already discussed how highly effective strategy users organise their thoughts and are flexible in their approaches to self-study, but is there any correlation with highly effective self-regulated learners, and their achievement level? McAndrew et al. (2016) reported high-achieving students are more likely to adopt effective self-study strategies based on robust research, in particular, self-testing, with low-achievers less likely to adopt self-testing as a revision strategy (Gurung, 2005). If this is the case, lower achieving students may benefit more from understanding the notions regarding

the advantages of the testing effect, as these students are less likely to adopt these strategies without assistance.

Oakley (2015) takes this further from the outlook of a student, and states that merely having the aspiration to learn, and spending time with material undertaking ineffective study will not be enough to succeed at mastery, and as a result. To this end, students who are metacognitively unaware of how best to use their self-study time, may result in large periods of valuable time under the illusion that they are studying, when in reality, no real change to lasting memory has taken place, and instead, 'labour in vain' revision takes control (Nelson and Leonesio, 1988).

However, the onus should not solely rest on the shoulders of the students to be responsible for making the correct self-study choices. Kornell and Bjork (2007) reported 80% of undergraduate students in their sample disclosed their choice of study strategy was improvised, revealing their teachers had failed to make an impact on the choices they made in unsupervised environments. Given the rigour of the new GCSE's and the higher volume of content the courses now comprise of, it would be unreasonable for students to navigate their own way through the educational landscape to discover best revision practice themselves. Inevitably this may lead to poor study choices at critical points in a students' education, and schools who do little in the way of educating their students of metacognitivley advanced self-study strategies should not be surprised if poor study habits develop.

2.7 – Self-regulation, grit and the creation of desirable difficulties

In order to create robust and lasting memories, and build skills and knowledge that will be transferable from the classroom into the exam hall, students must value and appreciate the benefits that testing as a learning strategy has to offer (Roediger et al, 2011), in addition to the diagnostic uses it may have (Amlund, Kardash, and Kulhavy, 1986). The advantages of engaging in such a learning strategy has been proven across a wide range of contexts, skills and abilities (Agarwal et al, 2012; Butler 2010), but are our students truly harnessing the power of retrieval practice to its full effect away from the classroom?

Bjork, Dunlosky and Kornell (2013) believes that to become more metacognitively advanced, learners must understand that the process is not straightforward, and will be required to go against the grain of instinct and familiar practice. With this in mind, Bjork (1994) developed the notion of desirable difficulties: the idea that learning is improved if the process is made more difficult. Bjork (1994) goes on to explain that should participants make retrieval practice more difficult for themselves, what they should experience is a depletion of short-term performance, but an enhancement in long-term retention. To students, this may sound unappealing, and a contributing factor into the intrinsic motivation they possesses may give insight into the likelihood that students will commit to creating desirable difficulties in unsupervised environments.

As a result, teachers must be cautious to strike a fine balance between creating difficulties in the learning process that will benefit students in the long-term, and making learning too difficult that they become disengaged. As Willingham (2009)

claims, if thinking is made to be too difficult, students will cease to think. It is therefore important for learners to experience a finely tuned equilibrium of desirable difficulties in order to enhance long-term memory (Bjork, 1994), with the necessary degree of success (Zimmerman, 2002) that will drive students to further engage in effective self-study without becoming disengaged by the by the learning process.

Self-regulated learning is described by Schunck and Zimmerman (1994) as the 'process whereby students activate and sustain cognitions, behaviours and affects, which are systematically oriented toward the attainment of their goals' (p.309). However, in order to sustain such behaviours, students must first become educated in effective self-regulatory strategies to be able to achieve such intrinsically set out goals. It is therefore important to emphasis to students what exactly constitutes effective self-regulatory strategies, and, perhaps more importantly, how to improve self-regulatory learning in less metacognitively advanced learners. Bandura (1997) defined metacognition as the awareness to ones thought process through mindful monitoring, evaluation and regulation of the cognitive process and being inherently aware of how best we learn in order to maximise performance. In other words, metacognition is thinking about thinking (Flavell, 1979; Schoenfeld, 1987). Stemming from Bandura's earlier phenomenon of self-efficacy, defined as 'a belief in one's capabilities to organise and execute the courses of action.' (Bandura, 1977, p.3), both processes are closely related to how well an individual is likely to perform on a task. Zimmerman (1995) concurs, and holds the opinion that motivation is crucial for self-regulated study.

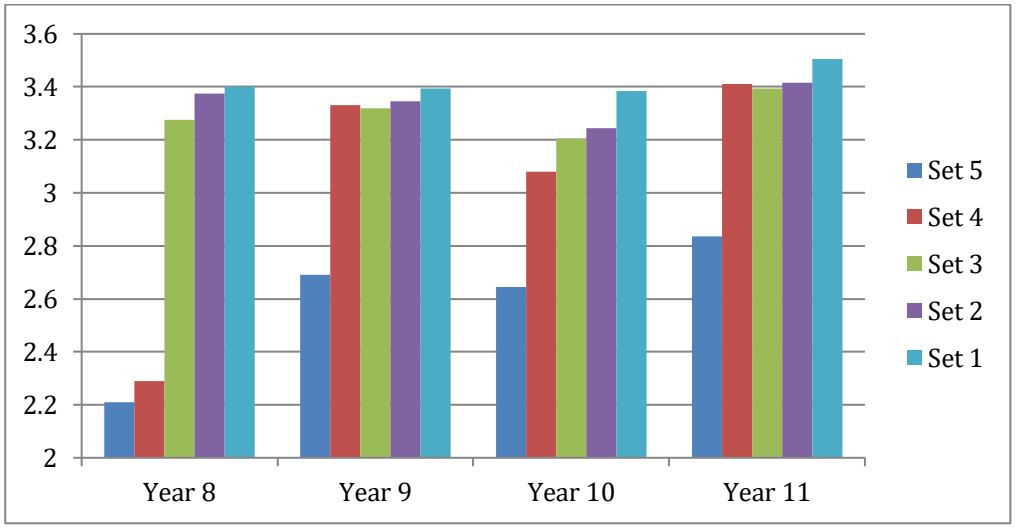
Inherent motivation is a notable contributing factor when examining self-efficacy, and numerous benefits have been observed when self-eficiacy levels are increased. Importantly, Lepper (1988) observed those who held larger internal amounts of self-efficacy, spend increased amounts of time on tasks, displayed durability when faced with problems, and, given an option, undertook problems that involved heightened levels of complexity. When taking into consideration Bjork's (1994) notion of desirable difficulties, it could be said that students who posses increased levels of intrinsic motivation are likely to engage more readily in self-testing behaviours.

Furthermore, a policy that has been engrained into practice for many years at my school is the concept of setting students by ability. Ability grouping suppresses the idea of an inclusive learning environment and calls attention to contrasting ability by dividing students into separate groups (Oakes, 1982). The EEF Teaching and Learning Toolkit (2013) are in agreement, and recommend that grouping students by their ability can 'undermine low-attainer's confidence and discourage the belief that attainment can be improved through effort' (p.6). Furthermore, Hattie (2011) suggest students placed in lower ability groups underrate their academic potential, which subsequently hinders their desire to progress. However, opposing arguments suggest that mixed ability groups, especially in older cohorts, is logical in terms of approach, where teachers are able to identify learning needs in relations to the gaps in knowledge their students may have (Ireson and Hallam, 2001).

With low levels of perseverance and lack of effort with regards to self-study observed in lower ability sets at my school, a look to literature on perseverance is useful. Grit, defined by Duckworth et al (2007) as the ability to persevere for the

benefit of long-term gain, can be measured using a simple questionnaire. As a result of the part two investigation for the MLT at Oxford University, it was discovered that lower ability sets consistently possess lower levels of grit. The results showed a uniformed response across year groups, with higher levels of grit reported in higher ability classes, whilst lower ability classes disclosed lower level of grit. With the year 10's at the time now in year 11 and about to embark on preparing for their GCSE examinations, could the revelation that lower grit scores, and therefore, levels of perseverance and effort at tasks, contribute to self-study choices that require less effort?

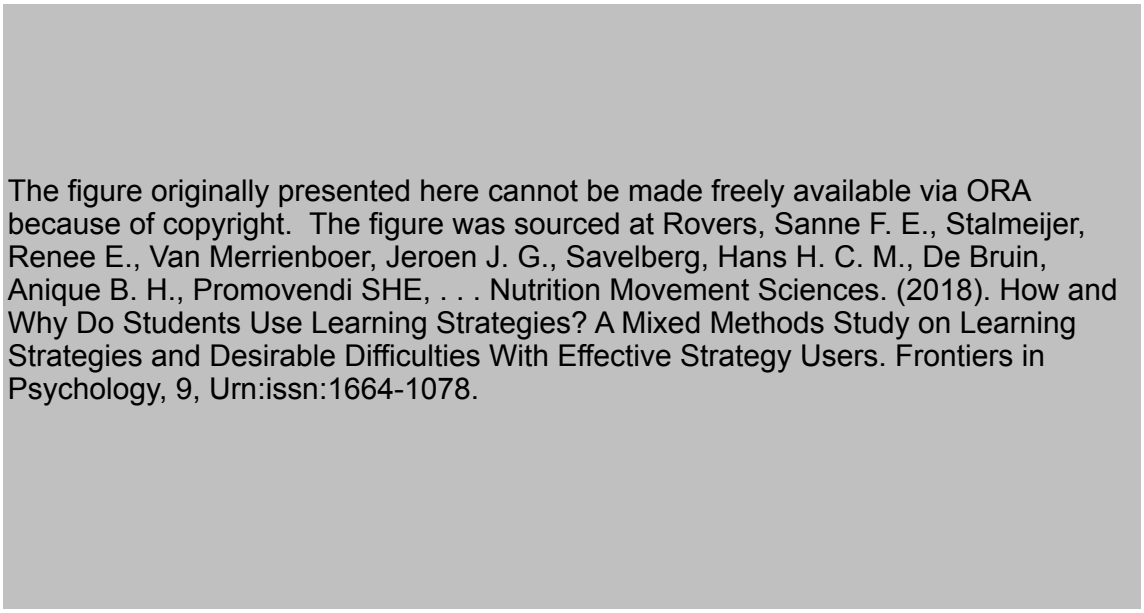
Figure 2.2 – Results of Grit questionnaire (Duckworth et al, 2007) from the MLT part two investigation



With, current teaching staff noting a culture of low self-esteem and lack of effort in lower ability sets, I am keen to discover any potential relationship between less effortful and therefore less effective study choices between higher and lower ability students.

In addition to the intrinsic motivational factors, an awareness of metacognitive knowledge would give promise to leading students towards effective self-study strategies. Metacognitive knowledge, a phrase developed by Flavell (1979), is defined as the knowledge that 'about what factors or variables act and interact in what ways to affect the course and outcome of cognitive enterprise' (p.907). Learners then put this knowledge into practice, where effective self-regulated learners are able to use an array of revision strategies to reach their learning targets. Additionally, Winne and Hadwin (1998) report that effective strategy users are aware that the quality of the revision strategy is directly related to educational performance.

A recent study exploring how highly effective self-regulated learners tackle self-study was conducted by Rovers et al (2018) is outlined in figure 2.3. Students' personal learning goals are the driving factor of the model, whereby they either opt for a quality or quantity attitude to their learning. Learners who favour a quantity approach will assign successful revision to the amount of work they have completed, whether it is effective or not. Rovers et al (2018) argues that the more successful of the two types of students are those who understand that it is the quality of the work being carried out, rather than the quantity, that is an essential quality of highly effective strategy users.



In favour of increasing academic performance, educating students to engage with self-regulatory strategies and habits has the ability to contribute to greater academic outcomes (Samson, 1985; Beidel, Turner and Taylor-Ferreira, 1999). But this is not to say one strategy is superior to any other. Although the many benefits of the testing effect have already been discussed, it is important to note how highly effective strategy users tackle self-study, based on academic research.

In contrast to some views whereby specific study strategies are cited as more beneficial than others for all students, (e.g. Tait and Entwistle, 1996; Sundre and Kitsantas, 2004), Rovers et al (2018) reported the use of less effective strategies by highly effective strategy users, but reported the reasoning behind this was students were able to adapt these less-effective revision strategies to meet the specific needs of their learning goals. A flexible approach to self-study based on the metacognitive knowledge that a student possesses is therefore desirable. Purdie and Hattie (1999) conducted a meta-analysis of 52 previously conducted studies to scrutinise the connection between effective self-study strategies and academic achievement. Where positive academic achievement was noted, the researchers concluded this was

not due to the innate quality of a specific self-study strategy, but rather, the presence of metacognitive understanding from the student. Having the ability to administer an extensive range of self-study strategies effectively irrespective of time, context and complexity of the material highlights the flexibility component of the model developed by Rovers et al (2018).

As a result of the demanding reform to the mathematics GCSE in 2017, increased pressure is placed on the ability of the student to effectively self-regulate their learning. With teachers experiencing time constraints due the increase in GCSE content, coupled with a lack of understanding of how memory works, and therefore, apprehensive to give up curriculum time in place of retrieval practice activities (Dunlosky et al, 2013), it comes as little surprise that an imbalance between the two competing priorities occurs. Consequently, students' perceptions of effective strategy choice play an unavoidably more prominent role, which could ultimately make an accumulation of novice learner decisions highly costly. Given that high levels of self-regulation and metacognitive awareness, along with a flexible approach to self-study are high predictors of academic success in a school based setting (Tait and Entwistle, 1996), it should become imperative to school leaders to find curriculum time in order to develop metacognitive knowledge (Flavel, 1979) and understanding so as to develop effective self-study strategy users (Rovers et al, 2018).

2.8 - Harnessing change

Having explored the possible reasons why all students may not deploy effective self-study within the realms of working memory capacity and metacognitive awareness, attention must turn to how we, as educators, might assist in encouraging students to change the habits of those who fall prey to ineffective self-study techniques through illusions of competency. Whilst Kornell and Bjork (2007) remind us that large proportions of student study technique is improvised, suggesting an accumulation of novice learner choices, should we be willing to direct students to reflect and change their habits if poor choices are being made?

The variations between accomplished self-regulated learners and novice self-regulators are that less metacognitively sophisticated students are incapable of changing their learning habits in response to the learning context presented (Zimmerman, 1998). But students will not entertain the idea of change if they are not motivated to do so. Psychology studies (Chartrand and Bargh, 1999; Wegner and Wheatley, 1999) draw attention to the idea that some behaviours, through repeated use and regular application, become instinctive. If regular practice of ineffective study habits have been ingrained in students' routines, it may become difficult for students, especially at a late stage in their studies, to be willing to alter bad habits. It is therefore imperative that teachers offer support and guidance in this area based on sound understanding of cognitive science and educational research.

That being said, the logistics and practicalities of developing student learning strategies can hinder the progress in this area. Students are not strengthening their learning toolkit and metacognitive awareness because schools are failing to designate

the time needed for students to reach a level of understanding that would warrant a change to their self-study behaviours (Hofer, Yu, and Pintrich, 1998). Recommendations from the Education Endowment Foundation (2018) recognise this, and state that time dedicated to an appropriate level of practice is an issue. They go on to suggest that schools that are looking to develop students metacognitive understanding should allocate time to teach students how best to learn so they can benefit from positive adjustments in academic behaviour. It is therefore not appropriate to simply tell students how best to learn and revise. Enough time must be spent exposing students to powerful learning tools that will warrant a shift in academic behaviours towards better self-study strategies.

Caution, however, must be taken. Students exposed to high levels of teacher instruction, whether it is teaching students what to learn, or indeed, how students learn best, would predictably experience high levels of success in supervised environments. The difficulty then becomes transferring the techniques demonstrated in controlled environments such as the classroom, to unsupervised self-study episodes (Clark, 1992). Success in the classroom does not necessarily translate to success in unsupervised environments, and overconfidence exhibited by students' illusions of competence (Glenberg, Wilkinson and Epstein, 1982) may distort their perceptions of effective self-study skills. Although some research suggests a demonstration of retrieval practice activity leads to increased levels of self-testing behaviour in unsupervised environments (Einstein et al 2012), the optimal way to achieve a change in study behaviour has yet to be discovered.

2.9 - Conclusions from the literature

Since the reform of the mathematics GCSE in 2017, and government influences in favour of a heavily knowledge based curriculum (Hirsch, 1987), added strain has been put on students not only to learn effectively in lesson time under the guidance of a teacher, but to effectively revise in unsupervised environments in preparation for rigorous exams. Student self-study choices are crucial, but do our students know how to revise, and do teachers know enough about techniques to support working memory to give advice driven by educational research?

The research explored in this literature review suggests that an imbalanced curriculum has seen teachers spend much longer teaching students what knowledge they need for high-stakes examinations, rather than dedicating time explicitly teach students how best to retain this critical information (Hofer, Yu, and Pintrich, 1998; Skemp, 1976; EEF, 2018). Furthermore, higher ability students, and typically, those with a higher working memory capacity, are more likely to engage with self-study strategies that are more effortful, but more fruitful to the long-term learning (Gurung, 2005), and as such, harness Bjorks's (1994) notion of desirable difficulties. Those who are classed as lower ability, have not only been found to engage less often with strategies that are less effortful and therefore, less effective, (Gurung, 2005) but are more likely to be unaware of the benefits testing and retrieval can have to learning (Einstein, Mullet and Harrison, 2012).

It is not, however, as straightforward as it seems. Although it reports that some strategies such as highlighting and rereading are ineffective, Rovers et al (2018) reported that highly effective strategy users often adopt flexible approach to their

studies, utilising both effective and ineffective techniques given the individual learning situation. This takes a degree of metacognitive awareness, which is inherently linked to the intrinsic motivation that Flavell (1979), speaks of. Students who are metacognitively more advanced, and understand key aspects of how they learn, are in an advantageous position when it comes to self-study choice (Zimmerman, 2002). Work by Duckworth et al (2007) takes this idea further, claiming that those who possess higher levels of grit and perseverance are more likely to achieve academic success. With strategies, such as self-testing requiring more effort, but producing lasting learning over time, Duckworth's (2007) notion of grit, coupled with Bjork's idea of desirable difficulties, interesting results conclusions may be drawn.

2.10 – Research questions

Having explored areas of research alluding to notions of effortful revision strategies, and those that are less effective, including the types of students who are more likely to adopt effective techniques in their self-study, I will attempt to answer the following questions:

- Does regular, low-stakes testing have an impact on student progress in high stakes examinations?
- Will regular, low-stakes testing increase self-testing behaviour as a study strategy?
- Will regular, low-stakes testing benefit higher and lower ability year 11 students, and if so, in what ways will it impact?

2.12 - Hypothesis

Given the literature surrounding self-testing behaviours, I hypothesise that in line with Gurung (2005) and McAndrew et al (2016), higher achieving students will disclose they use testing strategies more readily than lower achieving students. That being said, I believe that a demonstration of the testing effect during lesson time will impact both on higher and lower achieving students. Additionally, I hypothesise that the low-stakes test intervention will benefit both ability groups by contributing to enhanced progress from the previous year group, along with better organistaion of knowledge. A null hypothesis may see no significant change in progress between year groups, given the relatively far transfer of knowledge (Barnett and Ceci, 2002) and a sacrifice of learning time, and potentially disengagement from students who posses lower levels of grit and perseverance (Duckworth et al, 2007), given the more effortful nature of testing over other forms of revision (Bjork, 1994).

Chapter Three – Methodology

3.1 – Introduction

In this chapter, I will explain the methods adopted in the research project, comment on rationale for doing so, and ground these choices by evaluating their effectiveness. I will consider the option for choosing a practitioner based research model before going on to discuss the collaborative elements of my research project. I will close this chapter by considering the ethical implications of my research.

3.2 – Methodology

Practitioner research refers to 'forms of enquiry which people undertake in their own working contexts…to shed light on aspects of that work with a view to bringing

about some benevolent change.' (Dadds and Hart, 2001, p.41). Collaboration formed an integral element of the research project, and is fundamental to the personal and professional development of teachers and practitioners with Hargreaves (2000) pointing out that 'teachers normally learn better together than they do alone.' (p.165). As such, a practitioner based research framework was adopted for this project as it demonstrated qualities that both suited and complemented this type of method.

Given the nature of the research questions, in order attempt to obtain robust and meaningful conclusions, a mixed methods approach was chosen (Johnson, Onwuegbuzie and Turner, 2007) comprising of data gathered from questionnaires, student examination data, and interview responses from both students and teachers. This type of approach allows the triangulation of examination data and questionnaire responses gathered from students of varying abilities, with teacher and student views on self-study techniques, for a more rounded consideration of the issues (Denscombe, 2017). Furthermore, this type of data collection allowed members of the mathematics department to engage with practitioner research, with a view to gaining a more reflective approach to their own practice (Dadds and Hart, 2001).

The framework for the intervention was influenced by Hofer, Yu, and Pintrich, (1998) and recommendations from the EEF (2018), who believe it is not enough to just tell students about academically more robust revision strategies, but in order to observe changes in behaviour, sufficient time must be allocated to giving students adequate practice. It was therefore decided that students would be exposed to a demonstration of a retrieval practice based study strategy, with the aims being twofold; to expose students to a robust revision tool with a view to encouraging self-

testing behaviours in unsupervised environments, and, in addition to regularly engaging with retrieval practice, I hoped to assist in the transfer of skills and knowledge to the exam hall with a view to assisting increased attainment levels. After the delivery of an introductory presentation on the power of retrieval practice over time as a means of committing knowledge to long-term memory, in addition to highlighting what constitutes effective and ineffective study, students embarked on a study strategy demonstration. A weekly test comprising of 20 skills based questions was repeated over an 11-week window between year 11's November and February mock examinations, and a simple tracking system was used so he students were able to monitor their progress over the weeks.

Additionally, as part of my research analysis, the research questions aim to discover if learning has taken place, something that is difficult to capture genuinely. As such, a randomised control trial approach was adopted in order to analyse the differences between progress and viewpoints of a group of students who undertook an intervention.

3.4 – Intervention

3.4.1 – Outline of low-stakes tests

The practical hypothesis is that regular testing on chosen skills will increase retention of knowledge (Karpicke and Roediger, 2008). After a baseline questionnaire that aimed to discover students preferred strategy use, the low-stakes tests were created by the teachers of each year 11 class, to meet the needs of their students following their November mock examinations. 20 fluency based questions, which were repeated each week with slight tweaks to numbers or wording, were delivered to

students as a means of exposing them to a demonstration of how regular retrieval practice can improve learning over time (see appendix one).

Immediate feedback was issued to the students in the form of worked solutions the teacher had completed, and students marked their own once all students had completed all the questions. With the scores from the low-stakes tests not being analysed for the purposes of the project, data was drawn from other elements of the overarching intervention process to understand whether demonstrating the power of the testing effect to students would encourage testing behaviours as a self-study choice. Students recorded their own scores on a personal tracker, and were told for homework to work on an area that they could not answer correctly that week.

3.4.2 – Participants of study

The head of mathematics at the time the intervention was designed was keen for all students in year 11 to be exposed to the low-stakes tests, so as not to disadvantage any learners in the run up to their GCSE examinations. However, given the vast amount of data 258 students would generate, coupled with time constraints, it would be unwise for a single researcher to tackle such a large amount of data (Onwuegbuzie and Leech, 2007). This resulted in the decision to collect data from a typical higher and typical foundation ability set, diminishing the sample size to 54 (27 students in each class). Typicality was a theme I considered to be important, and as a result, decided to gather data from two contrasting ability classes; one higher and one foundation. The higher class had target grades ranging from 6-8 and the foundation class had target grades ranging from 3-5. Neither class had issues regarding behaviour, and both classes had a proportion of disadvantaged and SEN students that

were in line with whole school figures. Furthermore, the researcher taught the higher class, whilst the foundation class was taught by the second in department. Whilst it was important to consider the power imbalance that may have been created by the teacher researchers class, which is discuss further in section 3.6.3, I felt it necessary for both classes to have high standards of teaching. Additionally, in order to maximise consistency between the comparable year groups, it was requested at the start of the year to timetable the same teacher with the corresponding class as the cohort above in the previous academic year.

3.4.3 – Outline of collaboration

Although the collection of data remained with the two classes that had already been decided, I saw it as beneficial to the department if all staff were delivering the intervention, as views and opinions were also collected from all members of the department who delivered the low-stakes tests (seven members of staff plus the researcher) in the form of a semi-structured interview. So as to eradicate any subjective bias by having only two teachers involved in the delivery of the intervention, I saw it as an integral role in the triangulation of ideas and themes that occurred in a variety of different classrooms. Regular, informal discussions between teachers also took place in order to share experiences of carrying out the intervention.

3.4.4 – Introduction to the project

An introductory lesson on effective self-study was delivered to all year 11 classes (see appendix eight), which outlined why students would be doing weekly, low-stakes tests. The two classes of students, whose data from questionnaires and focus group interviews, were told at this stage that the low-stakes tests formed part of a research project.

3.5 – Data sources

3.5.1 – Questionnaires

Two, key data collection points were fundamental to the research project, and it was at these points that the questionnaires were used in order to extract student views and opinions on their self-study choices. Questionnaires were distributed both before the intervention (see appendix two) and after the intervention had taken place (see appendix three), as a means of understanding whether their opinions and views had shifted towards more metacognitively robust self-study strategies.

The decision to use questionnaires was one rooted in their nature to collect a considerable amount of useful data in a relatively short amount of time, in addition to this being the traditional way of determining the self-study strategies of students (Panadero, Klug and Järvelä, 2016). Research in the field of student self-study widely makes use of questionnaires as a means of capturing self-study behaviours (e.g. Entwistle and McCune, 2004; Pintrich, Smith, Garcia, and McKeachie, 1993). However, these studies did not specifically seek out to assess whether self-testing behaviours were being adopted. As such, the questionnaire format was influenced by Karpicke et al (2009), who specifically sought to determine whether students

participate in retrieval practice in preparation for exams.

Students in this study were given an open response option to list their commonly used strategies, and then rank the strategy they used the most frequently. I believed it to be important for this study not to replicate much of the research questionnaires produced in previous investigations, which give predetermined lists of strategies for students to rank either their effectiveness or their use (Kornell and Bjork, 2007), as this has been found to restrict the authenticity of the responses of the students (Schuman and Presser, 1996). It was anticipated for the purposes of this investigation that the most common forms of self-study students would report using when revising for mathematics would by practice papers, practicing recall (any other form of testing), using flashcards, highlighting and underlining, rewriting notes and rereading. Any other forms of self-study mentioned would be reported, but within the parameters of this research, would not be discussed at length.

Additional questions not included in the research of Karpicke et al (2009) were incorporated into the post-intervention questionnaire in order to draw out why changes to study technique as a result of the demonstration intervention may or may not have occurred. Furthermore, I wanted to determine whether students were willing to incorporate the retrieval tests into their self-study given the option.

Furthermore, as a result of the motivational factors which shed light on reasons to why students may or may not be adopting for effortful, yet more effective self-study techniques, I saw it necessary for students in year 11 to reassess their grit scores (see

appendix 9). I anticipated that these may have changed from the previous year given the added pressures of high-stakes mock examinations.

In keeping with the same reporting system of Karpicke et al (2009), and in light of relatively small numbers of students in each ability group (each class had 27 students), percentages were used to report initial self-study strategy, and any subsequent percentage changes as a result of the demonstration intervention.

3.5.2 – Semi-structured interviews – Students

Two small focus groups comprising of two higher and two foundation students were selected to take part in semi-structured, group interviews, both before the intervention and after. A total of four interviews took place with four participants. The students were selected based on a combination of their initial questionnaire results and differing opinions of self-study, and the relationship with the researcher. I saw it important to the project to select students to be interviewed who I knew well from either previous or current teaching experiences, and would be able to give coherent, genuine and relaxed responses to the interview questions.

The interviews were used as a means of shedding more light onto the responses from the questionnaires, assisting in more robust conclusions from the research project. Group interviews were decided on as this format is one whereby participants usually feel more at ease, and therefore, more likely to share their opinions with the interviewer (Foddy, 1994). Opportunities are also presented whereby participants may agree, remark or disagree on each other's comments as the framework presents

a potentially less intimidating atmosphere as opposed to a one to one interview technique.

Semi-structured interviews provided the most suitable framework, which enabled me to ask the most applicable questions about both the low-stakes tests, how students prepare for exams, and whether the tests had impacted on subsequent self-study technique decisions. Additionally, this structure allowed the interview to deviate to reflect a normal flow of conversation, and touch on areas that were not previously planned for. In particular, this qualitative form of data collection allows for students to describe what strategies they may adopt for particular learning goals through elaborate explanations. It has also ben suggested by Rovers et al (2018) that by interviewing students, it may be possible to differentiate between the value students give to a particular study strategy against their actual practical use.

All interviews were recorded, and students were made aware of this before each of the interviews. Recordings were checked on the day to ensure the quality was suitable to transcribe. Partial transcription was decided as a means to reduce the time it took to transcribe the interviews (Altrichter, Posch and Somekh, 1993), whilst gaining a rounded yet accurate insight into each participant's opinions. This allowed the researcher to be available to respond to any comments or ask further questions based on what the students had said, rather than writing up notes after each response. All student interviews were completed during the usual school day, with registration time in the morning being utilised for this.

3.5.3 – Semi-structured interviews – Teacher

Many of the reasons for deciding to carry out semi-structured interviews for students, apply to the teacher interviews also. The interview was held after the intervention had taken place in order to reflect on how successful teachers felt the demonstration had been, and how responsive and engaged students had been throughout the process. Eight teachers who had all engaged in delivering the low-stakes test intervention to year 11 students took part in the interview. It was believed to be beneficial for teachers to be interviewed as a collective where opinions could be shared and comments could be made to answers from other participants. A semi-structured approach was again taken in order for the researcher to gain authentic views from the perspective of the teacher.

3.5.4 – Mock examination data

With the purpose of the intervention to encourage and nurture the use of self-testing behaviours in student's revision time, it was important to assess the degree to which engaging with additional acts of retrieval practice impacted on high-stakes examinations. Mock examination data was used therefore used to compare the progress students made between their November and February in the selected classes who were exposed to the intervention, with the corresponding classes in the cohort in the year before. The class teachers did not mark their own students' mock examination papers, so as to remove any bias with regards to the impact on the research project. Nonetheless, this is regular practice for the mathematics department, as a means of avoiding inconsistencies in teacher marking. Grade boundaries were used for the year that the past paper was sat, so as to keep in line with national progress.

It would be useful for this part of the discussion to explain exactly what and how progress data is measured against both nationally and at my school. Along with the introduction of the newly reformed GCSE specification in Mathematics and English in 2017, with remaining subjects following in 2018 and 2019, the introduction of an alternative way of measuring student progress came into effect. Instead of the headline figures of how many students achieved A* - C, the new school accountability measure introduced in 2016, now measures how much progress each child makes from entering secondary school. Progress 8 figures now drive schools to focus on any child underachieving in any subject, rather than placing a heavy emphasis of intervention on previous C/D borderline students. To calculate a progress 8 score, each grade that a student achieves in any subject at GCSE level will be awarded a set amount of points. Numerical points on the new 9-1 grading system are equivalent to the grade achieved, commonly known as an attainment 8 score. Taking the attainment 8 score for each student and subtracting the estimated attainment 8 score for that subject calculates the progress 8 scores for individual subjects. Government recommendations confirmed that progress 8 scores would pivot around 0, with the majority of schools reporting progress 8 figures of between -1 and 1. A school scoring a progress 8 figure of 0 will translate to students performing just as well, on average, as those with comparable prior attainment on a national level.

3.6 – Ethical considerations

The University of Oxford's Research Ethics Committee (CUREC) granted approval for the research to be carried out in November 2018. Permission from the head

teacher of the school was also given after an outline of the research to be carried out was communicated through a short meeting (see appendix 13). Additionally, throughout my research project, the Ethical Guidelines for Educational Research (BERA, 2011) recommendations were adhered to.

3.6.1 – Confidentiality

The school the project was carried out at will remain anonymous throughout this research report. Both student and teacher participants of the research project were assured that confidentiality would remain throughout the process. All participants were told they would remain anonymous in any reporting of data.

3.6.2 – Consent

Permission from parents of student participants of the research project was not obtained, as procedures used for the collection of data from these students were deemed as low-risk. Additionally, the intervention design was not deemed to be outside normal classroom practice. As such, this fell in line with the Modus Operandi set out in the MSc Learning and Teaching guidelines.

Students who were in the chosen two classes for data collection were briefed regarding the additional tasks (questionnaires for all and interviews for some students) that would be required of them, and voluntary informed consent was obtained verbally from all participants (Thomas, 2009). All students in both classes gave this consent.

Students who were selected to take part in the small, group interviews were given the choice of whether to opt in or out. All four students gave consent. Additionally, members of staff involved in the research project were asked for their verbal consent once the research had been outlined to them.

3.6.3 – Power imbalance

When the researcher teacher interviewed students, it is important to recognise the power imbalance situation that could occur (Cohen and Manion, 1980). It was crucial for the students being interviewed to be able to express both open and frank opinions without the feeling of providing wrong answers (Aubusson et al, 2009). Furthermore, the possibility that the teacher-researcher has to influence any opinions must be minimised in order for any responses to be able to be scrutinised critically (Pring, 2000).

In order to resolve such issues, I opted to gather additional evidence from a neutral third party source through the use of questionnaires (Matthews et al, 2002). Furthermore, I took the decision to remain as the interviewer of the focus groups so I could intervene and guide the interviews as the researcher, and as a result, guide conversation to areas of specific interest (Altricher et al, 1993).

Before each interview that took place, students were reassured by the researcher that the aim was not to give correct answers, as there weren't any. Rather, the researcher was interested in finding out about what pupils thought of self-study and the intervention that took place. Additionally, there may also have been a degree of power imbalance experienced by the other teachers taking part in the research

project. Given my positioning within the department at the time of designing and implementing the intervention in collaboration with the Head of Department, other members of staff may have experienced a top down approach to the project. In order to alleviate the possible power imbalance issues, all members of staff were explicitly told that their involvement in the research project would have no impact or influence on any performance management related targets.

3.6.4 – Protecting data

Al data collected as part of the research project, which included audio recordings of staff and student interviews, were kept in secure conditions. Any electronic files were stored on a password protected USB memory stick, and persons used for the research project remained unidentifiable throughout.

3.6.5 – Avoiding disadvantage

When the intervention design was put forward to the head of mathematics at the time, she felt it important that no pupils were disadvantaged by not being exposed to the powers of the testing effect. With the literature suggesting large educational benefits in relation to the use of testing as a means to assisting memory and retention of knowledge, it was decided that all year 11 pupils would be exposed to the weekly low-stakes tests so as not to disadvantage any student. However, data was only collected from two pre-determined classes in order to gather data to shed light onto whether the intervention had an impact on study preference and behaviours.

3.6.6 – Low-stakes tests

The low-stakes tests lessons took the full 60 minutes, which included a flexible amount of time for students to complete the tests, with the additional feedback from the teacher. Tests took place on a weekly basis for a total of 11 weeks. There is scope here to argue that the 11 hours used for the intervention were taken from usual learning lessons, and could be classed as a considerable amount of time given the total contact hours allocated to year 11 over the academic year (around 100 hours), and therefore contributing to depriving students of time to explore what Skemp (1976) describes as relational understanding. However, given the gravitas of the literature surrounding the testing effect in the assistance of long term memory and retention of knowledge, coupled with little time my school currently allocated in assisting students with metacognitive strategies to assist with self-study, it was decided by the Head of Mathematics at the time that this would be deemed as a useful way of using year 11 contact time for the whole cohort, so as not to deprive any student of taking part in a strategy that has such positive effects, whether it would increase self-testing behaviours as a revision strategy or not.

Students were also informed that the scores from the low-stakes tests would not be collected. This meant that students were marking their own work and recording their scores on a tracker sheet, without the teacher checking work. Some students could have misconstrued this as an opportunity to become negligent towards the process, and therefore disengage with the demonstration, thus, compromising on their learning.

Chapter four – Findings and discussion

4.1 - Introduction

In this chapter I will discuss the findings of my research project first by outlining the key discoveries in each section with relation to the literature reviewed, before exploring these in more detail through evidence previously discussed. I will then attempt to answer my research questions in relation to the literature I have explored in chapter two by outlining essential points at the start of each section, before discussing why these findings may have surfaced.

4.2 - Curriculum content

- Some teachers felt one hour per week was the limit of what time they could devote to engaging with retrieval practice, as there was too much content to cover.
- Students thought curriculum content for mathematics is heavy, and they were under a large amount of pressure to memorise techniques, formulae and perform a variety of skills.

All teachers in the department agreed that the low-stakes tests were a good use of lesson time, and that their students, whether they had made an impact on self-study choices or not, had benefitted from engaging in regular retrieval practice. *'It's almost a different mentality for learning',* one teacher remarked. *'You have to get away from thinking that doing bog standard lessons is the way kids learn. You spoon feed them more than you realise when you do that.'*

This research has not only highlighted the importance of exposing teachers to methods grounded in educational and cognitive research will induce advice of a higher quality when educating students in revision techniques (Dunloskey et al, 2013). Edging away from the mind-set that performance is the same as learning (Soderstrom and Bjork 2015), and the notion that familiarity by spoon-feeding can contribute to an illusion of competence (Jacoby et al, 1994), could a shift in teacher attitude pave the way for stitching retrieval opportunities into an abundance of learning opportunities?

What may be difficult, however, would be convincing teachers that time dedicated to retrieval practice would not be time wasted, and in line with Hofer, Yu, and Pintrich, (1998) teachers commented on their hesitation to sacrifice curriculum time to teach student how to learn rather than what to learn. With one teacher commenting, *'There is just too much to cover',* and others agreeing that since the GCSE reform in 2017 where substantially more content would have to be covered in the same amount of time, this may cause teachers to feel added pressure.

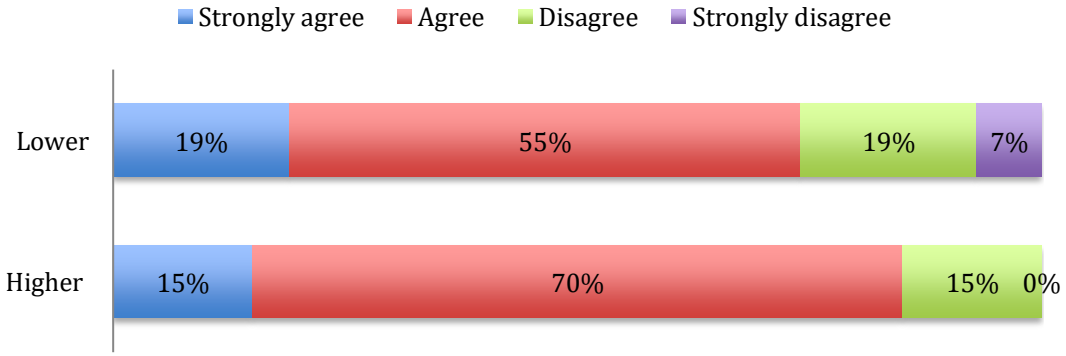
Another teacher agreed, and suggested the weekly retrieval demonstration was the *'absolute limit'* of the time she could commit regarding her allocation of curriculum time to engaging with retrieval practice. With Skemp (1976) reminding us that teachers are split between teaching what to learn and how to learn it, difficulty if met when attempting to gain balance in this domain. Seemingly, in addition to convincing students of the benefits of the testing effect, teachers also needed to be persuaded to detach from old stigmas that testing is used as a tool for assessing intelligence (Black and Wiliam, 1998), and instead, being more intellectual as

practitioners in the dedication of curriculum time to develop metacognitive understanding in our learners.

The views of the students united with those of the teachers. When asked what teachers do to help with revision, both ability groups of students were of the opinion that little time is spent teaching students how to revise. The views of Skemp (1976) resonated with responses from all of the interviewed students, with one maintaining *'there is just so much stuff to learn'*. In keeping with the teacher opinions regarding the imbalance time allocation, it appears the decision to prioritise new learning over metacognitive understanding is putting learners at a disadvantage. By not allowing students to learn from experts, in this case, teachers, it would be unreasonable to assume that students would become anything more than novices when it comes to effectively deploying self-study techniques.

That being said, responses to the statement 'I found the tests an effective use of lesson time' drew parallels between the two ability groups:

Figure 4.1 – Higher and lower ability responses to the question 'The tests were a good use of lesson time'



The majority of both classes (85% of the higher ability and 74% of the lower ability) responded positively towards the testing demonstration, and voiced the opinion that using lesson time to engage with a revision demonstration was a good use of lesson time.

With advice from (Flavel, 1979) supporting the notion that school leaders should allocate necessary curriculum time in order to develop student's metacognitive knowledge, and further recommendations from EEF (2018) suggesting metacognitive awareness could be improved by dedicating contact time to developing students' learning behaviours, is it now time to break old habits and look beyond devoting large amounts of the curriculum to content rather than strategy use?

4.3 – The testing effect

- Both higher and lower ability students, even after the low-stakes test intervention, were unaware of the potent benefits regular testing can have, and viewed the weekly tests largely as a tool for diagnosing knowledge.

A similarity between student responses to what they liked about the low-stakes tests came as an interesting surprise. Despite an introductory presentation outlining the benefits of repeated testing to improve retrieval strength, rather than solely using tests for diagnostic purposes, it was seemingly not enough to change students' perceptions of using testing to truly harness their many other benefits (Karpicke and Roediger, 2008). One foundation student remarked, *'there were too many of the same question'*. This observation alludes to findings from Karpicke and Roediger

(2008), suggesting that students are unaware that repeated retrievals strengthens lasting memories (Wheeler and Roediger, 1992), and the nature of the tests were to repeatedly revisit specific mathematical skills.

Although a comment from a higher student suggested a more positive interaction with the intervention, a similar theme occurred in that students remained unaware of the many learning advantages retrieval practice has.

> *'They really helped me put topics into categories of what I know and what I don't know. If I knew I didn't get something one week, I'd revise it so I knew how to do it the week after.'*

The views of this student resonate with opinions from (Einstein et al, 2012), in that the important advantages that have been mentioned are those of a diagnostic nature, rather than to harness the retrieval and retention rewards testing can present (Wheeler and Roediger, 1992). Despite this, by identifying gaps in her knowledge (Amlund, Kardash, and Kulhavy, 1986), the student has identified that better understanding of metacognitive monitoring will be gained (Kornell and Son, 2009) by organising her knowledge into learnt and unlearnt material.

When probed further about what the student meant specifically by *'revise'*, she suggested that she would *'read over'* her notes before the low-stakes test the following week. Although this is by no means a fruitless way of making use of testing, perhaps more could be done to educate students about the vast amount of advantages self-testing revision strategies hold in order to shift away from making use of ineffective habits. That being said, what the higher student demonstrated here

is what Rovers et al (2018) described as the flexible element when deploying strategy use. The student was able to adapt her strategy use depending on the situation, rather than solely relying on 'labour in vain' strategies (Nelson and Leonesio, 1988).

4.4 - Transfer of knowledge

- Progress data gathered from the November and February mock examinations revealed the higher ability group gained 0.17 grades per student, which indicated progress was -0.15 grades per student lower than the previous year group.
- In contrast, students in the lower ability group gained, on average, one grade per student between the months of November and February. This indicates an increased rate of progress from the previous year, with the current group exceeding their progress by an average of 0.72 grades per person between February and November.

Low-stakes testing can assist in helping students to practice specific and isolated skills, and supports committing said knowledge to long-tem memory through repeat attempts thus, making it more difficult to forget (Roediger et al, 2011). However, findings from my research project have revealed that the benefits of low-stakes testing are complex when attempting to measure the impact they have had on high-stakes examinations, and have exposed differences between the ability groups. Whilst it was expected that students made progress in the tests over the weeks in an isolated, near transfer setting (Barnett and Ceci, 2002), many students raised concern over the transfer of knowledge to a high-stakes setting. It would be unwise to assume that successful near transfer of knowledge should lead to successful far

transfer of the same knowledge. In some cases, in particular with higher ability students, this has not been the case, contrasting with Karpicke and Aue's (2015) argument that complex material can be transferred to far settings.

The table below outlines the progress 8 figures obtained by the higher and lower ability classes from 2018, and the equivalent classes in 2019, with the difference in progress between the two-year groups calculated.

Figure 4.2 – Comparison of progress data for higher and lower ability group classes

	Class of 2018		**Class of 2019**	
	Higher	Foundation	Higher	Foundation
November P8 figure	-0.45	-1.39	0.21	-1.82
February P8 figure	-0.13	-1.11	0.38	-0.82
Overall progress	0.32	0.28	0.17	1.00
P8 Difference			-0.15	0.72

Between the November and February mock examinations, students in the higher class gained on average, per student, 0.17 of a grade. In contrast to this, the results in the foundation level class saw students gaining on average 1 grade per student, between the months of November and February. Students in the lower ability class appear to have successfully transferred knowledge from the low-stakes test setting into the exam hall, more so than their higher ability peers. This indicates a significant difference between the progress of higher and lower ability students, considering some students in the foundation class had remained on a static grade since the beginning of their GCSE course two years prior.

Censabella and Noël, (2008) shed light on the differences between students who possess higher working memory capacities, and note that these students are in an educationally more advantageous position to begin with. It may be the case that it is the lower ability students who have benefitted more from the intervention as a result of engaging with a process that aims to increase working memory capacity, given that they were already lagging behind their peers in terms of working memory capacity (Swanson and Jerman, 2006).

However, attributing the progress to the low-stakes test intervention is difficult, and a number of other factors may have influenced the progress for the lower achieving students, or lack of for their higher ability peers. An alternative reason to attribute differences in progress data may be the differences in difficulty between the two examination series. However, steps are taken to minimise this effect due to the variations in grade boundaries between year groups, which directly reflects the difficulty of the exam.

4.5 - Initial study techniques

- Higher ability students reported using self-testing strategies more than lower ability students, which is in support of my initial hypothesis.
- Despite claims that rereading, rewriting and highlighting are ineffective ways to study, both higher and lower ability students are seemingly unaware of this, with both groups of students reporting heavy use of rereading and rewriting.

- Lower ability students reported a less flexible approach to their studies, and associated quantity, regardless of quality of learning, to success.

Notably, the pre-intervention results for the allocation of time students spend engaging with particular study techniques, revealed that lower ability students did not engage with self-testing techniques as much as their higher ability peers.

Figure 4.3 – Results of baseline strategy use, and strategy used the most often

Strategy	% of students who listed strategy		% who ranked as #1 strategy	
	Lower	Higher	Lower	Higher
Past papers	85%	70%	26%	59%
Flashcards	63%	41%	11%	11%
Practice recall	30%	78%	0%	22%
Rereading	67%	48%	33%	7%
Rewrite notes	70%	56%	30%	0%
Highlighting	19%	7%	0%	0%
Other	7%	11%	0%	0%

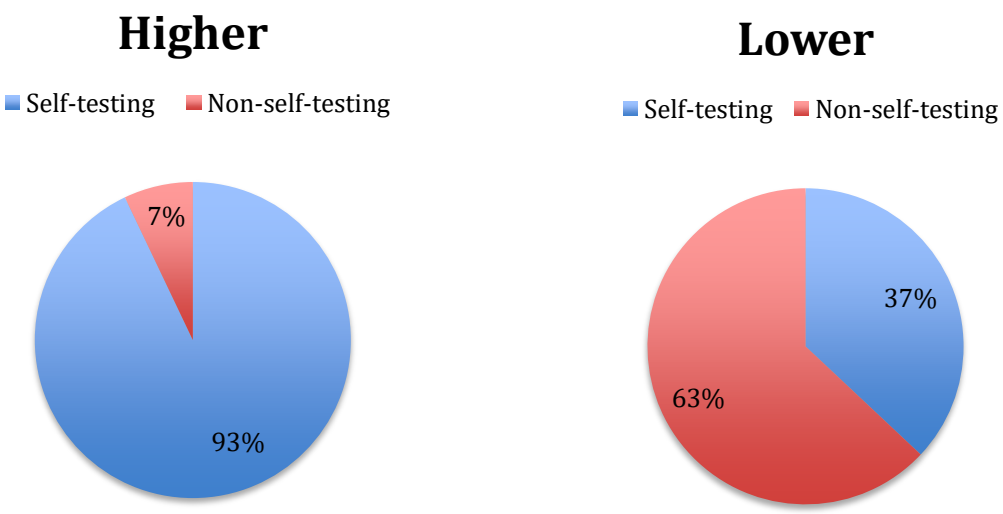
A promising baseline set of strategy use for higher ability students revealed that past paper practice and recall were the groups two most commonly used methods for study. That being said, there was still room for improvement, as despite 59% of higher pupils revealing completing past papers was their highest ranked study technique, only 11% reported favouring flashcards, and 22% favoured practicing recall. Furthermore, 48% and 56% of higher ability students reported using rereading and rewriting respectively as a study strategy, despite claims that these strategies are ineffective. However, no students ranked rewriting or highlighting as their first choice of study method, with only 7% of students ranking rereading as their top choice. Although this baseline assessment of study strategies for higher

ability students paints a promising picture, with many students favouring highly effective study strategies, it not only pinpoints initial significant differences between the two ability groups strategy choice, but also highlights further room for improvement would be advantageous.

The lower ability student's baseline assessment of preferred study strategies presented a contrasting picture to their higher ability peers. Despite past paper practice appearing as the strategy used the most by these students, with 85% reporting they make use of the technique, use of less effective strategies are reported by high numbers of the students. Rewriting and rereading are techniques used by 70% and 67% of the lower ability students respectively, with rereading appearing as the most popular first ranked strategy. No students listed practicing recall as their preferred technique. Rewriting and rereading came above past papers as a number one ranked strategy.

When compared alongside each other, clear differences are seen when strategies are classed as 'self-testing' or 'non-self-testing'. 'Other' strategies were excluded.

Figure 4.4 – Pre-intervention comparison of strategy time allocation when classed as 'self-testing' vs 'non-self-testing'



Despite many students in the higher ability class reporting self-testing strategies as their preferred methods of study-technique, it was still deemed necessary for the intervention to be carried out with both ability groups, given that many students in both ability groups reported using techniques that have been found to be ineffective. However, the extent and the reasoning behind this ineffective use of self-study was one that was difficult to capture through use of a questionnaire. It was therefore important for the research to gain additional perspectives from the students by conducting small group interviews, in order to gain a more rounded perspective of why particular revision techniques are favoured over others. In order to delve deeper into the reasons behind students use of ineffective revision technique, influence was taken from Rovers et al (2018) mixed methods study, whereby interviews following a baseline questionnaire determined how chosen revision strategies were put to use.

Higher ability focus group interviews revealed some interesting initial findings. In keeping with the literature and the baseline questionnaire results, it was noted that students of a higher ability both had an appreciation for self-testing revision strategies, and actively engaged with these when studying in unsupervised environments. When asked why schools test students, one student commented, *'As much as it is useful for the school, it is also useful for us to know where we are at'*. Again, the notion that tests are solely used for diagnostic purposes is highlighted here, giving more weight to the strong sense of culture that tests are used to assess intelligence, as opposed to viewing them as a learning device (Einstein et al, 2012).

It was deemed important to query student's thoughts and opinions on less effective revision strategies that they reported using as revision techniques. When asked about

copying notes from either a textbook or a student's workbook, a higher achieving student commented on creating her own desirable difficulties.

> *'I think it is a useful method sometimes, but I would never just copy out the notes word for word. I don't see the point in that. I would turn it into my own words because that way you feel like you are learning more.'*

This demonstrates what Rovers et al (2018) described as the 'active processing' stage of their model of highly effective strategy users learning, whereby effective strategy users are continually creating desirable difficulties for themselves in order to maximise their learning potential. Additionally, students who go against the grain of what they may intuitively believe about revision strategies is a step towards becoming metacognitively more advanced (Bjork and Bjork, 2011).

Lower achieving students held a conflicting view to the higher achieving students with regards to rewriting notes, which provided further stimulus for the design of the project. One student remarked, '*I'll copy out the notes from a revision book so I have my own revision page that I can file away',* whilst another remarked, *'I have loads of notes. It helps the information go in when I copy it out.'* This duplication of material coupled with the notion of a quantity approach rather than recognising the quality of the revision taking place (Rovers et al, 2018), indicates a strong sense that lower achieving students require guidance and support to assist them in utilising and adopting more effective self-study strategies that favour quality over quantity.

Baseline data gathered from the students to determine what strategies they use and how effective they deem them to be revealed that higher achieving students actively engage more with self-testing study strategies than their lower achieving peers which resonates with Hartwig and Dunlosky's (2012) findings. Despite some higher

achievers, and a larger proportion of lower achievers confessing use of less effective self-study strategies, it was the higher ability students who reported a more flexible approach towards strategy use, whereas lower achievers associated quantity with success, regardless of the quality of the learning that was taking place (Rovers et al, 2018).

4.6 - Self-regulation, grit and desirable difficulties

- Teachers reported that lower ability students took longer to engage with the intervention as they found them initially difficult, whilst higher ability teachers noted students feeling motivated by seeing their scores increase.
- Teachers of lower ability classes observed lack of effort and engagement in any revision for some students. This could be linked to findings from the part two assignment where levels of grit were lower in lower ability groups.

A difference was observed when teachers were asked about student engagement with the low-stakes tests in lesson time. Teachers of higher ability classes noted high levels of engagement, with one teacher stating, '*they like seeing their results go up every week. I think its very satisfying for them to see'*. Teachers of the lower ability group observed a contrasting attitude, noting that it took a long time for the students' marks to begin to improve. A lower ability teacher commented:

> *"They were not sold for a long time because they didn't improve. If they weren't improving then I think they just thought what's the point? They got there in the end, but it took a lot of encouragement from me to get them on board with it."*

The satisfaction that the higher ability teacher speaks of in terms of student experience was noted much earlier on in the 11-week process than the lower ability students. As Willingham (2009) claimed, if thinking is made to be too difficult for students, then learners will not only cease to think, but levels of motivation, in particular in unsupervised environments, may be affected. As this intervention was primarily set up to convince less effective strategy users to adopt self-testing behaviours more readily, it may have become counter productive if the students became disengaged by the low-stakes tests, therefore, making them less appealing to adopt as a revision technique. It may have been the case that in this instance, the balance of desirable difficulties (Bjork, 1994) and necessary degree of success (Zimmerman, 2002), may not have quite been in equilibrium for the lower achieving students.

One lower ability teacher commented of the revision strategies her students adopt at home, stated:

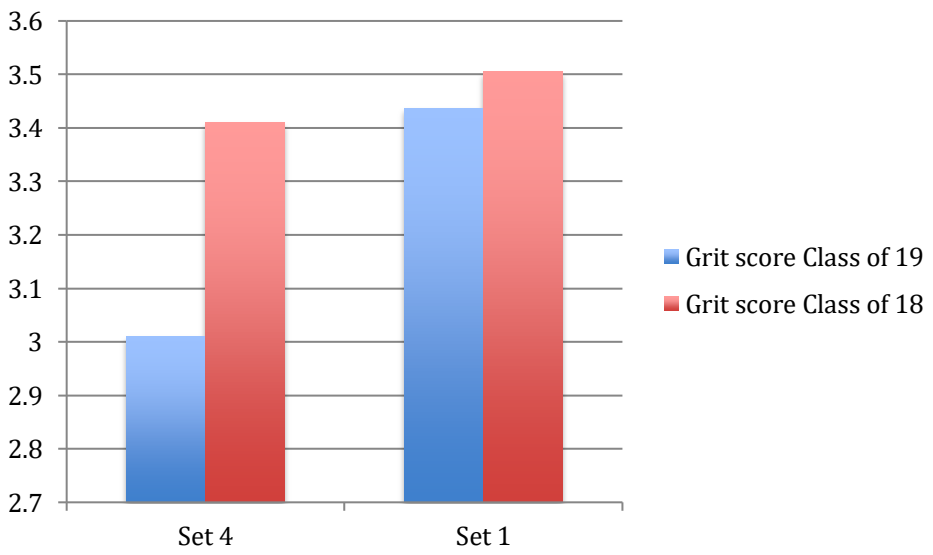
> *'I genuinely don't think some of these kids ever do things like this at home. They might do the odd past paper, but few and far between. They always take the easy way out.'*

This comment supports the students' responses to the study technique questionnaire, whereby lower ability students reported using less effective strategies than higher ability students. If learners do not regularly engage in retrieval practice techniques in unsupervised environments, and teachers have reported little in the way of regular use if retrieval techniques in lesson time, the low-stakes tests may have contributed to a limited amount of retrieval based activity in some student's repertoire of study. With the literature suggesting that low achievers are less likely to adopt self-testing

as a revision strategy (Gurung, 2005), and the results from out data collection in line with this, the tests may have served as a major increase in the amount of retrieval practice that some students engaged with.

Dunlosky et al (2013) recorded reasons for students not taking up effective self-study techniques in unsupervised environments was the small amount of effort required to partake in actions such as rereading or highlighting. The increased amounts of effort displayed by both higher students in their responses draws attention to the findings from the part two research. Increased levels of grit have previously been noted in the highest ability groupings. With heightened levels of grit comes a greater desire to persevere at a task by exerting sizeable amounts of effort (Duckworth et al, 2007). Increased engagement with more effortful techniques could therefore be a result of enhanced levels of grit, and as a result, an appreciation and awareness that effective learning may take longer than techniques that have the capacity to distort a student's judgement of learning through what Jacoby et al, (1994) describes as illusions competence.

Figure 4.5 – Average class grit questionnaire results for higher and lower ability groups in 2018 and 2019



Results from the Grit questionnaire show that higher ability sets in both cohorts displaying a higher average grit score than their lower ability peers. As such, both classes in the higher ability set are more likely to persevere at more effortful tasks (Duckworth, 2007), which gives more understanding as to why lower ability students may not engage in revision strategies that require more effort and are therefore more effective for long-term learning.

Agarwal et al (2017) states that testing as a learning device can serve as a tool to 'level the playing field' (p.770) between lower and higher ability students. With students in lower sets traditionally making less progress than higher ability classes at my school, this finding could serve as a pivotal reason to make use of the testing effect, in order to benefit students who may possess less metacognitive awareness (Shaughnessy and Zechmeister, 1992; Kornell and Son, 2009).

It must be noted here the large difference in grit score for the lower ability sets in each cohort. Given the lower levels of effort and perseverance that the 2019 lower ability group reported, it appears that the low stakes test intervention has been essential to exposing students with more effortful learning and revision techniques. If students are less likely to engage in retrieval practice at home, encouragement and engagement in lesson time may be a way of leading students towards effective study-strategies.

4.7 – Harnessing change

- The biggest change in study habit was observed in the increase of reported practicing recall from higher achieving students.
- Decreases in ineffective study habits were observed with rereading for lower ability students and rewriting for higher ability students.
- 63% of higher ability but only 32% of lower ability said they would use the tests as a revision resource if they had access to them.

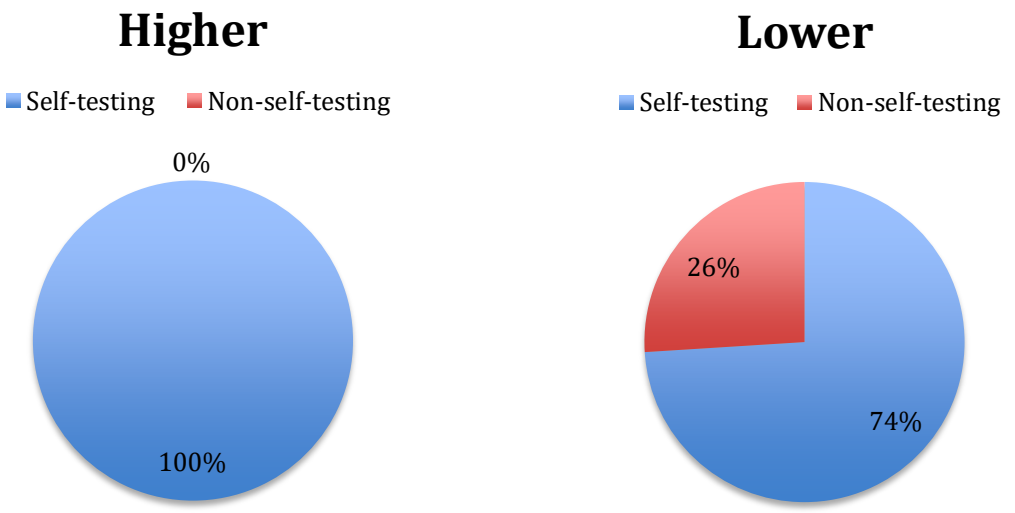
It became apparent from the baseline questionnaire results that students were overwhelmingly keen to change and improve their self-study strategies if the opportunity presented itself. 68% of higher students and 83% of foundation students said they would appreciate guidance and assistance with effective revision strategies, and would be willing to change their methods if they had both the time and resources to do so. Seemingly, on the surface, a large proportion of students express their desire to improve their metacognitive knowledge in order to better understand how best to self-study (Flavell, 1979). It is perhaps a shame that this issue had not been recognised or addressed earlier.

Figure 4.6 – Percentage change of strategy use, and strategy used the most often

Strategy	% change of students who listed strategy		% change who ranked as #1 strategy	
	Lower	Higher	Lower	Higher
Past papers	+7%	0%	+30%	+4%
Flashcards	-7%	-7%	-4%	-7%
Practice recall	+44%	+19%	+11%	+11%
Rereading	-22%	-11%	-11%	-7%
Rewrite notes	-11%	-7%	-26%	0%
Highlighting	-37%	-7%	0%	0%
Other	0%	0%	0%	0%

When viewed alongside each other, not only can we still see some difference in time allocation of self-testing strategies, but also changes after the intervention are now clear to see.

Figure 4.7 – Post intervention comparison of strategy time allocation when classed as 'self-testing' vs 'non-self-testing'



Pleasingly, all higher ability students reported the strategy they allocate the most time to was classed as a self-testing technique. It must be mentioned, however, that higher students already reported high levels of self-testing strategy use. The largest changes were observed in the strategy use that lower ability students ranked as their preferred technique, with a 37% increase favouring self-testing over non-self-testing methods. A promising admission from the students given teachers reported reluctance in the opening weeks of the demonstration.

In essence, students reportedly did increase self-testing behaviours as a result of the intervention, more notably the higher achievers, who were keen to incorporate the intervention resources into their self-study time in the run up to their GCSE exams in

the summer term, in addition to many students reporting dropping ineffective forms of self-study from their repertoire of techniques.

Seemingly, with one higher ability boy, the low-stakes test intervention made him reflect on his strategy use at home:

> *'It's made me realise how important it is to practice questions yourself. You can do things copy out your notes and things like that, but you can't then just be expected to do it yourself. You need to concentrate yourself on being able to do it.'*

The higher ability boy's view on individual concentration in order to ensure understanding resonates with McAndrew et al (2016) who reported higher achieving students are more likely to adopt effective self-study strategies than their lower achieving peers. This view was also consistent with a higher achieving girl:

> *"They have made me do more practice questions. Before, I would just read through my exercise book and that's all I would do, but now I would do the actual questions, I will go and find them online where you tell us to and just practice doing the questions."*

There is no question that the intervention has had an impact on both lower and higher ability students, albeit in differing ways. Although reluctance was observed from teachers and demonstrated by lower ability students, engaging with the intervention demonstration appears to have triggered not only an increase in progress data, but an exposure to strategy use that some students may rarely have engaged with at all outside the classroom.

With the majority of students in Kornell and Bjork (2007) study suggesting their revision strategies were self-taught, questionnaire responses from both groups of students were in keeping with this literature, with 93% of foundation and 78% of higher students reporting improvised self-study techniques. This resonated into the post-intervention interview responses. One higher student commented *'Nobody really teaches you this stuff, so you're just left to like, figure it out yourself, and I don't really know what I'm doing to be honest.'* This concurred with a comment from a lower ability student, stating, *'I've just taught myself as I've gone along.'*

If most students are reporting improvised revision strategies, along with many learners, particularly those who fall in the lower ability category, reporting use of ineffective revision techniques, are we as practitioners letting our students down by not actively seeking to explicitly teach students about the benefits of robust revision techniques?

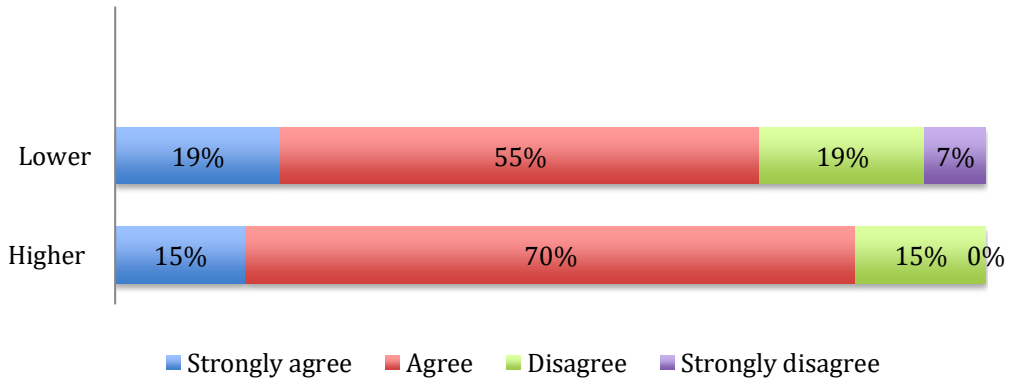
In line with recent recommendations from EEF (2018), both higher and lower achieving students identified a need to actively engaging in a revision strategy in lesson time to demonstrate to them how it worked would more likely make them change their choice of strategy in the future. In turn, this sympathises with the views of Hofer, Yu, and Pintrich, (1998), who advocate allowing time for students to develop metacognitively advanced techniques, rather than leave students to develop without explicit teacher guidance.

In light of this, it transpired that some students became aware that some of their techniques were ineffective, and communicated an appreciation for the intervention.

One lower ability student remarked, *'I wish we could have done something like this sooner so I could have used it all the way through school.'* She, along with the other student present in the low ability interview, expressed difficulty in finding resources, and as a result, would gravitate towards ineffective study techniques where the action requires no more material other than a textbook such as rereading or rewriting.

A discrepancy was, however, observed. Despite lower ability students reporting on the post-intervention questionnaire that their self-testing habits had increased, they were less inclined to use the tests at home if they were to be made available (52% disagreed). Although it was by no means essential that students made use of the resource away from the classroom, given the opinion of one student commenting that she struggled to find resources of a high quality, it seems that when given the opportunity, lower ability students are still reluctant to engage with more effortful self-testing revision.

Figure 4.8 - Higher and lower ability responses to the question 'If the tests were made available, I would use them at home as part of my revision'.



A closer look at the reasons students had assigned to the decision not to use the tests as part of their revision revealed yet further interesting differences between higher and lower ability students. Of the higher ability students who revealed they would not use the intervention tests as a resource for their own revision, three out of the four students voiced that they preferred to use practice papers instead. Given using practice papers is still a form of self-testing, this reasoning does not present concern. However, of the 7 lower ability students who expressed they would not use the tests as part of their revision communicated a conflicting opinion. They reasoned that the tests were *'boring'*, *'repetitive'*, and *'difficult'*, with one student remarking that he preferred to read his notes rather than to test himself, as this was *'easier'*

With research from Bjork (1994) stating that desirable difficulties are favourable to the self-study process, and the notion that repeated retrievals through testing can assist in long-term retention of knowledge (Ewers and Buonanno, 2003), it may be the case that lower achieving students have misinterpreted the nature of the intervention. Despite knowledge of their increase in grades following the intervention, students were of the opinion that engaging in weekly retrieval practice episodes had little to do with their improved exam results, and failed to look past initial feelings of boredom and having to work hard to achieve success. These findings resonate with McCabe (2011), who suggest that students are unaware of the lasting benefits retrieval practice brings to learning.

Chapter 5 - Conclusions

5.1 – Conclusions

The aim of this research project was to explore the self-study choices adopted by higher and lower achieving students when preparing for high-stakes mathematics examinations, and to investigate whether self-testing behavior as a study-strategy increased after a demonstration of the testing effect through a series of weekly, low-stakes tests. The grounding of this research project was the necessity to guide less metacognitively sophisticated learners to adopt effective learning strategies more often, as opposed to gravitating towards less effortful, and therefore less effecting techniques, so as to improve their revision experience.

Research question 1: Does regular, low-stakes testing have an impact on student progress in high stakes examinations?

When compared to the previous years progress 8 data, a clear and positive difference between the lower ability students is observed. This complements the responses of the lower ability to the post intervention questionnaire, which suggested that they had adopted more self-testing strategies in their study time than their higher ability peers, which supports my hypothesis when applied to lower achieving students. With higher ability students already reporting high levels of self-testing before the intervention, but dropping less effective study strategies after the intervention it may be the case that lower ability students have seen added benefits to their progress data because they have adopted self-testing strategies more than higher-ability students. It cannot, however, be deemed as conclusive that the significant progress of the students in the lower ability class is as a result of the low-stakes test intervention. Many other factors may have contributed to the increased progress, such as

motivational differences, and social backgrounds, but the intervention has addressed important issues to improve year 11 self-study strategies at the request of both students and teachers. The hypothesis that the low-stakes tests intervention has made an impact on the progress of higher ability students is not supported, as students of this ability made slightly less progress than the cohort who did not receive the intervention. With higher ability student's baseline strategy use reporting more frequent use of self-testing strategies than the lower ability groups, it could be the case that lower ability students have benefitted more as a result of the intervention.

Research question 2: Will regular, low-stakes testing increase self-testing behavior as a study strategy?

Both higher and lower ability students revealed that their revision now incorporated more self-testing strategies as a result of the intervention, supporting the hypothesis that the testing intervention increased self-testing behaviours for both ability groups. However, a discrepancy was observed with lower ability students revealing that if they had the choice, the majority would not use the tests as a way of revising at home. This is in light of some lower ability students revealing they struggle to find resources that make use of self-testing, which they attributed as a reason for gravitating towards less effective strategies such as rereading. One has to question if lower ability students were reporting an increase in self-testing behavior at home, but in reality, reverting back to old habits in unsupervised environments. A larger proportion of the higher ability students (63%) requested copies of the intervention to use as a revision tool at home, reporting that they found them a 'useful' way to build on weak areas of their subject knowledge.

Research question 3: Will regular, low-stakes testing benefit higher and lower ability year 11 students, and if so, in what ways will it impact?

Before the low-stakes test intervention took place, many students disclosed that they were unaware of the benefits testing can have to long-term memory. What was surprising was that even after the intervention, a large number of higher and lower ability students reported they saw the tests useful to diagnose learnt and unlearnt knowledge, rather than appreciating testing as a way of improving memory.

With lower ability students detailing an increase in self-testing behavior, but open responses to the intervention largely negative, commenting on the increased effort engaging in such a strategy would take, the intervention served as a surprising support tool. With the part two investigation revealing a correlation between ability set and perceived level of effort and perseverance at a task, lower ability students have seemingly benefitted from the support of a supervised revision strategy demonstration, as many students, particularly those in the lower ability set, reported using fruitless techniques often in their revision time. Despite whether lower ability students increased their self-testing behavious in unsupervised environments or not, it appears to have increased their progress whilst simultaneously engaging students with retrieval practice activity, which some students did little in the way of outside the classroom.

School leaders must take inspiration from literature encompassing long-term learning, and apply such advice to their school setting by investing time, rather than spending ample amounts of money on short, quick fix, solutions that inevitably end up having little to no lasting impact on long-term retention of knowledge.

5.2 – Limitations of the project

5.2.1 – Time

Perhaps the most obvious limitation to the research project is that of time. Given the nature of repeated retrieval practice with the idea of incorporating time to forget as part of the lasting learning process (Bjork and Bjork, 1992), a longer period than was accessible within the realms of this research project would have been more beneficial to the conclusiveness of the progress data presented in the mock examinations.

5.2.2 – Questionnaires

Although questionnaires are a useful research tool to enable a substantial amount of data to be gathered in a short space of time, limiting factors such as interpretation of language, genuine responses and unambiguous answers from students contribute to difficulties in gaining authentic perspectives and opinions.

Additionally, another difficulty encountered through the use of questionnaires to gather data was being able to understand whether the responses given by students were genuine or not. Providing responses through questionnaire surveys is deemed the most appropriate technique for assessing study technique according to Bjork et al (2013), but there is no guarantee that students self-report reflects genuine responses (Delamont, 2002), making it difficult to understand behaviours in unsupervised environments. Furthermore, there is scope for students to affiliate their perceived value, as opposed to their authentic use, to a particular technique (Bråten and Samuelstuen, 2007; Bernacki et al., 2012).

Given that the class teachers of the sample classes used had already been teaching the students for at least one whole academic year previous to the research project commencing, teachers had already built secure relationships with the students. As a result, some members of staff commented on the fact that there may have been elements of teacher-pleasing with regards to responses to the questionnaires, and students were possibly attempting to report what the teacher or researcher wanted to hear, rather than communicating a genuine response.

5.2.3 – Sample size and population

Notably, the unavoidable differences between the cohort of students who received the intervention with comparison the cohort in the year above meant that factors such as teaching style, reflection and advancement of teaching practice despite corresponding classes having the same teacher, and attitudes of the whole cohort, limited the authenticity of the findings and could therefore not be deemed to be conclusive.

The issue was raised with the Head of Department, and discussions were held regarding whether to carry the project forward with a younger year group, which would alleviate the factor of exam pressure and time constraints. However, it was deemed that the year 11 cohort was in need of an intervention not only to assist retention of knowledge and key skills, but one which would encourage better self-study behaviours outside of the classroom. Additionally, an older cohort, particularly year 11, would have an increased pressure to revise for high-stakes examinations given they would be sitting them at the end of the academic year: a pressure that is impossible to replicate in a younger year group.

5.3 – Implications for practice

As a result of this research project, a number of implications have come to light in response to the research questions posed:

- The senior leadership team implemented a whole school timetabling strategy whereby the week before any year group sit one of their tri-annual examinations, lessons in all subjects are spent dedicated to teaching and demonstrating to students how to revise the relevant material effectively. This accumulates three weeks of teaching, equivalent to 12 hours of contact time in each core subject now solely dedicated to showing students how to learn, rather than exclusively focusing on what to learn.

- End of topic assessment in GCSE mathematics, which occur fortnightly, now dedicate one third of the marks to previously taught material.

- Training, including attendance at external professional development courses and internal guest speakers, have been attended by all members of the mathematics department, focusing on differences between learning and performance, balancing curriculum time to support long-term retention, and low-stakes testing implementation in the classroom.

- Low-stakes tests have since been added to all year groups' schemes of work, with a particular emphasis on year 7 and 8 in order to embed the idea that testing is one of the most effective revision strategies from a much earlier age.

- The low-stakes tests now happen more frequently, but contain fewer questions and therefore take less time, so as not to occupy a whole lesson or disengage students, particularly those of a lower ability.

5.4 - Closing remarks

This study has not only taken a departmental focus of wanting to improve revision techniques among year 11 students, but has generated wider implications far beyond what was initially set out to achieve. With whole school strategies now in place to develop self-regulated learning techniques across all departments in the school, it has been humbling to extend the collaboration that started within the mathematics department to a school wide audience.

Yet whilst the ripples of the research project continue, I find it important to note the implications for my own practice. For the students I teach, daily retrieval practice is now the norm, and the proportion of time devoted to remembering rather than performing has grown. The research that I have undertaken has not only allowed me to fully appreciate and understand the difficulties some students face when the time comes to revise, but has also taught me to empathise, encourage and support students to optimise their study habits for better academic success.

Appendix 1: Example of low-stakes test

1. $1\frac{2}{3} \times 1\frac{1}{4}$

2. Using $x_{n+1} = -3 - \frac{2}{x_n^2}$

with $x_0 = -3.5$

(a) find the values of x_1, x_2 and x_3

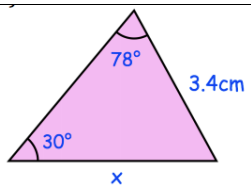
3.

Year 9	Year 10	Year 11
100	50	50

The table shows information about the number of students in a year group. Work out how many from each year group if a stratified sample is to be taken.

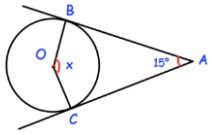
4. The base of a triangle is 30cm, correct to 2 significant figures. The height of the triangle is 40cm, correct to 1 significant figure. Calculate the upper bound for the area of the triangle

5.

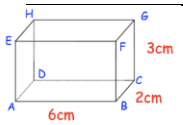


6. The number of days, D, to complete research is inversely proportional to the number of researchers, R, who are working. The research takes 125 days to complete when 24 people work on it. Find out how many people are needed to complete the research in 60 days.

7.

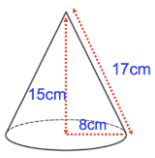


8. Work out AG.



9. Solve: $8x^2 - 8x - 9 = 0$

10. Surface area:

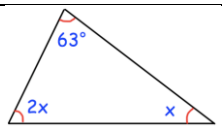


11. $\mathbf{a} = \begin{pmatrix} -5 \\ p \end{pmatrix}$ $\mathbf{b} = \begin{pmatrix} q \\ 1 \end{pmatrix}$ Given $\mathbf{a} + \mathbf{b} = \begin{pmatrix} 1 \\ -7 \end{pmatrix}$ work out p and q.

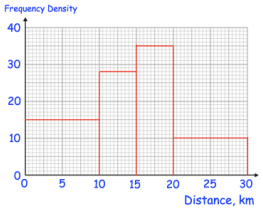
12. Evie is given a 22% pay rise. Her new salary is £21960 What was Evie's salary before the pay rise?

13. Mrs Martin wants to open a new restaurant in her town. She wants to find out what type of food people in her town like. Design a suitable question she could ask to find out what type of food people like.

14. Find x.



15. Estimate how many students live between 10Km and 14Km away from the college.



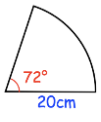
16. $\frac{2x-1}{9} + \frac{x+2}{3} = 0$

17. Write down the equation of a circle that has centre (2, 7) and has a radius of 6.

18. (a) Complete the table of values for $y = 3^x$

x	-2	-1	0	1	2
y					

19. Arc length?



20. Solve the following:
$x^2 + 8x - 10 = 0$

Appendix 2: Pre-intervention student questionnaire

1). List the strategies you regularly use to revise for mathematics exams:

……..

……..

……..

……..

……..

……..

……..

……..

2). Which strategy do you use the most often?

……

3). Who taught you how to revise? (Please circle)

- A teacher

- You are self-taught

- A friend

- A family member

- Other (Please state) ……..

3). I know how to revise properly for maths exams:

Yes: ☐ No: ☐ Not sure: ☐

4). If I knew of a better way to revise that would improve my grades, I would use it:

Yes: ☐ No: ☐ Not sure: ☐

Appendix 3: Post-intervention student questionnaire

1). List the strategies you regularly use to revise for mathematics exams:

……..

……..

……..

……..

……..

……..

……..

……..

2). Which strategy do you use the most often?

……

3). The tests were a good use of lesson time:

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree

4). If the tests were made available, I would use them at home as part of my revision:

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree

5). Why or why would you not use the tests to help with your revision at home?

.....

.....

.....

.....

.....

Appendix 4: Lower ability pre-intervention study strategy questionnaire

Student	#1 Ranked strategy	Other listed strategies		
A	Rewriting	Past papers	Flashcards	Rereading
B	Past papers	Rereading	Practice recall	Flashcards
C	Flashcards	Past papers	Highlighting	Practice recall
D	Rewriting	Flashcards	Past papers	Rereading
E	Rereading	Rewriting	Past papers	Other
F	Rereading	Rewriting	Highlighting	Other
G	Rewriting	Rereading	Flashcards	Past papers
H	Rereading	Highlighting	Past papers	Rewriting
I	Flashcards	Past papers	Practice recall	
J	Past papers	Flashcards	Rereading	Highlighting
K	Rereading	Past papers	Highlighting	Rewriting
L	Past papers	Flashcards	Practice recall	Highlighting
M	Past papers	Practice recall	Rereading	Flashcards
N	Rereading	Rewriting	Flashcards	Highlighting
O	Rewriting	Past papers	Rereading	Highlighting
P	Rereading	Flashcards	Past papers	Rewriting
Q	Past papers	Practice recall	Flashcards	
R	Rereading	Highlighting	Past papers	Rewriting
S	Rewriting	Highlighting	Rewriting	Flashcards
T	Past papers	Rewriting	Practice recall	Highlighting
U	Rewriting	Rereading	Rewriting	Highlighting
V	Flashcards	Past papers	Highlighting	
W	Rewriting	Past papers	Flashcards	
X	Rewriting	Highlighting	Practice recall	Past papers
Y	Past papers	Rereading	Flashcards	
Z	Rereading	Past papers	Rewriting	
AA	Rereading	Flashcards	Past papers	Highlighting

Appendix 5: Higher ability pre-intervention study strategy questionnaire

Student	#1 Ranked strategy	Other listed strategies		
A	Practice recall	Rereading	Rewriting	
B	Past papers	Practice recall		Other
C	Practice recall	Flashcards	Practice recall	
D	Past papers	Flashcards	Practice recall	
E	Past papers	Practice recall	Rereading	
F	Past papers	Rereading	Flashcards	Practice recall
G	Past papers	Rereading	Rewriting	
H	Past papers	Rewriting	Highlighting	
I	Past papers	Practice recall	Rewriting	
J	Past papers	Past papers	Practice recall	
K	Flashcards	Rewriting	Rereading	
L	Past papers	Flashcards	Rereading	
M	Flashcards	Practice recall	Past papers	
N	Practice recall	Rewriting		
O	Practice recall	Rewriting	Rereading	
P	Flashcards	Past papers	Practice recall	
Q	Rereading	Rewriting	Highlighting	Other
R	Past papers	Practice recall		
S	Practice recall	Rewriting	Flashcards	
T	Practice recall	Rereading	Rewriting	
U	Past papers	Practice recall	Rereading	
V	Rereading	Flashcards	Rewriting	
W	Past papers	Practice recall	Rewriting	Other
X	Past papers	Rereading	Practice recall	
Y	Past papers	Rewriting	Flashcards	
Z	Past papers	Rewriting	Rereading	Practice recall
AA	Past papers	Flashcards	Practice recall	Rewriting

Appendix 6: Lower ability post-intervention study strategy questionnaire

Student	#1 Ranked strategy	Other listed strategies		
A	Past papers	Practice recall	Rewriting	Other
B	Past papers	Rereading	Practice recall	Flashcards
C	Flashcards	Past papers	Highlighting	Practice recall
D	Past papers	Flashcards	Practice recall	
E	Past papers	Rewriting	Practice recall	Other
F	Past papers	Rewriting	Highlighting	Practice recall
G	Past papers	Rereading	Flashcards	Practice recall
H	Rereading	Highlighting	Past papers	Rewriting
I	Practice recall	Past papers	Rewriting	
J	Past papers	Flashcards	Rereading	Highlighting
K	Practice recall	Past papers	Highlighting	Rewriting
L	Past papers	Flashcards	Practice recall	Highlighting
M	Past papers	Practice recall	Rewriting	Flashcards
N	Rereading	Rewriting	Flashcards	Highlighting
O	Past papers	Practice recall	Rereading	Highlighting
P	Rereading	Flashcards	Past papers	Rewriting
Q	Past papers	Practice recall	Flashcards	Practice recall
R	Rereading	Highlighting	Past papers	Rewriting
S	Past papers	Highlighting	Rewriting	Practice recall
T	Past papers	Rewriting	Practice recall	Highlighting
U	Rewriting	Rereading	Rewriting	Highlighting
V	Flashcards	Past papers	Highlighting	Rewriting
W	Practice recall	Past papers	Flashcards	Practice recall
X	Past papers	Highlighting	Practice recall	
Y	Past papers	Rereading	Flashcards	Practice recall
Z	Rereading	Past papers	Rewriting	Practice recall
AA	Rereading	Flashcards	Past papers	Highlighting

Appendix 7: Higher ability post-intervention study strategy questionnaire

Student	#1 Ranked strategy	Other listed strategies		
A	Past papers	Rereading	Practice recall	
B	Past papers	Practice recall	Other	
C	Practice recall	Flashcards	Practice recall	
D	Past papers	Flashcards	Practice recall	
E	Past papers	Practice recall	Rereading	
F	Past papers	Practice recall	Flashcards	
G	Past papers	Rewriting		
H	Past papers	Rewriting		
I	Past papers	Practice recall		
J	Past papers	Past papers	Practice recall	
K	Practice recall	Rewriting	Rereading	
L	Past papers	Flashcards	Rereading	
M	Flashcards	Practice recall	Past papers	
N	Practice recall	Rewriting		
O	Practice recall	Rewriting		
P	Practice recall	Past papers	Practice recall	
Q	Practice recall	Rereading	Other	
R	Past papers	Practice recall		
S	Practice recall	Rewriting	Flashcards	
T	Practice recall	Rereading	Rewriting	
U	Past papers	Practice recall	Rereading	
V	Past papers	Flashcards	Rewriting	Rereading
W	Past papers	Practice recall	Rewriting	Other
X	Past papers	Rereading	Practice recall	
Y	Past papers	Practice recall	Flashcards	
Z	Past papers	Practice recall	Rereading	
AA	Past papers	Flashcards	Practice recall	

Appendix 8: Powerpoint for introducing the intervention to students

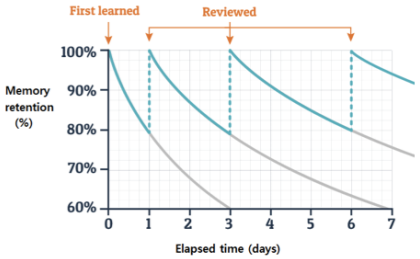
Revision Techniques In Mathematics



Your memory is key to success, so this year we will be doing lots of things to help you to REMEMBER what you learn in lesson time.

Retrieval practice is bringing something to mind that you have been taught in the past. The more times you attempt to remember something, the better it will stick in your long-term memory.





Each time we learn something new, we want to try to remember it as many times as possible so we don't forget it.

When you revise at home, some strategies have been proven to be better than others:

Past papers

Rewriting notes

Highlighting

Flashcards

Making review sheets

Testing yourself

Underlining

Testing yourself has lots of benefits when it comes to revision. We want to demonstrate these to you.



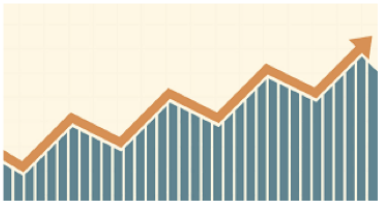
Each week you will do a test, but these tests don't mean anything.
They are just a tool for you to learn!

Because of this, I won't be collecting in your scores.
They don't matter ☺

Sheet	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mark														

Question	Topic	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10	Test 11	Test 12	Test 13	Test 14
1	Fraction arithmetic														
2	Iteration														
3	Sampling														
4	Upper and lower bounds														
5	Sine rule														
6	Direct and Inverse Proportion														
7	Circles Theorems														
8	3D Pythagoras and Trigonometry														
9	Solving quadratics														
10	Cones and spheres														
11	Vectors														
12	Percentages														
13	Questionnaires														
14	Equations and Shapes														
15	Histograms														
16	Equations and problems														
17	Equation of a circle														
18	Growth and Decay														
19	Arc Length and Sector Area														
20	Completing the square														

You will be able to track your scores for each question, so you can work on areas you are struggling on.



The tests will contain the same topics for a number of weeks. You should see your scores improve. If you don't, make sure you work on a topic you got wrong for the following week.

The tests will contain the same topics for a number of weeks. You should see your scores improve. If you don't, make sure you work on a topic you got wrong for the following week.

We are doing this to demonstrate to you how powerful testing yourself can be.



After your mocks in February, we want you to tell us if they have worked for you, and whether you use more testing when you revise at home.

Appendix 9: Grit questionnaire

	Not like me at all	Not much like me	Somewhat like me	Mostly like me	Very much like me
1. New ideas and projects sometimes distract me from previous ones	5	4	3	2	1
2. Setbacks don't discourage me. I don't give up easily.	1	2	3	4	5
3. I often set a goal but often choose to pursue a different one.	5	4	3	2	1
4. I am a hard worker.	1	2	3	4	5
5. I have difficulty maintaining my focus on projects that take more than a few months to complete.	5	4	3	2	1
6. I finish whatever I begin.	1	2	3	4	5
7. My interests change from year to year.	5	4	3	2	1
8. I am diligent. I never give up.	1	2	3	4	5
9. I have been obsessed by a certain project or idea for a short time, but later lost interest.	5	4	3	2	1
10. I have overcome setbacks to conquer an important challenge.	1	2	3	4	5

Appendix 10: Questions for student interview pre-intervention

1. How do you know if you have learnt something?

2. Who has taught you how to revise?

3. What do you think makes a good revision strategy?

4. How do you know if the strategies you currently use work?

5. How often do your teachers spend advising you how to revise?

6. Would it be beneficial to you if someone were to teach of a better way to revise?

7. If you knew of a better way to revise, would you be willing to change you strategies?

8. What do you do if you find revision difficult?

9. What would be the best way to teach you how to revise effectively?

10. Is there anything that your teachers do that doesn't help you to revise?

Appendix 11: Questions for student interview post intervention

1. Did you find the tests impact on your learning?

2. If the tests impacted on your learning, in what ways did they do so?

3. Were you able to transfer the knowledge you recalled in the tests to your mock exams?

4. Did you enjoy the tests?

5. Which elements of the tests did you enjoy?

6. Have the tests had any impact on the way you revise at home?

7. As a result of the tests, have you changed the way you revise for math's exams?

8. If the school were to teach students how to revise for their exams again next year, how could we make it better?

Appendix 12: Questions for semi-structured teacher interview

1. How often do you take the time to teach students how to revise for exams?

2. How well have the students engaged with the testing intervention?

3. If students engaged well/not so well, why do you think this is?

4. Previous to the intervention, how well do you think your students were revising at home?

5. Do you feel like your students have benefitted from taking part in the intervention? If so, in what ways?

6. If we were to teach pupils how to revise in the future, what would be the best way to go about this?

Appendix 13: CUREC ethical approval form

CENTRAL UNIVERSITY RESEARCH ETHICS COMMITTEE (CUREC)

Form CUREC 1A Checklist for the Social Sciences and Humanities



The University of Oxford places a high value on the knowledge, expertise, and integrity of its members and their ability to conduct research to high standards of scholarship and ethics. The research ethics clearance procedures have been established to ensure that the University is meeting its obligations as a responsible institution.

They start from the presumption that all members of the University will take their responsibilities and obligations seriously and will ensure that their research involving human participants is conducted according to the established principles and good practice in their fields and in accordance, where appropriate, with legal requirements. Since the requirements of research ethics review will vary from field to field and from project to project, the University accepts that different guidelines and procedures will be appropriate.

- Please check **"Where and how to apply for ethical review"** and the **CUREC flowchart** first to see if you need ethics approval.
- Please complete this form using a word processor and email it, together with your supporting documents, to your **Departmental Research Ethics Committee (DREC)** (**if applicable**). If you don't have a DREC please email this form to **ethics@socsci.ox.ac.uk** using your official **ox.ac.uk email** address. **Only emailed applications will be accepted.**

WHAT THIS CHECKLIST IS DESIGNED FOR

This **CUREC 1A checklist** is designed largely for research that falls within the Divisions of Social Sciences and Humanities where ethical issues are relatively few and straightforward. Interviews, field work and oral history are also included in the CUREC process.

The **full CUREC 2 application** is only required where certain project characteristics (e.g. type of participants, or procedures) result in a more complex set of ethical issues. It is expected that only in a limited number of cases will it be necessary for researchers to complete a CUREC 2 application. The checklist below will direct you to a CUREC 2 application if needed.

WHAT THIS CHECKLIST WILL NOT ASSESS

This checklist does not cover research governance, satisfactory methodology, or compliance with the requirements of publishers when administering their tests or questionnaires. As **principal researcher (i.e. principal investigator)**, it is your responsibility to ensure that requirements in these areas are met.

CUREC does not review studies classed as **audit** (see Glossary and Decision Flowchart for CUREC on our website).

If your study involves **NHS patients, NHS staff / data / facilities, or human tissue**, please check the Decision Flowchart for NHS approval and contact the Clinical Trials and Research Governance (CTRG) team in the first instance.

Further information on the University's research ethics procedures is available from the CUREC website.

SECTION A: Filter for CUREC2 application

This section determines whether your study raises more complex issues which require the completion of a full application for ethical review, known as the CUREC 2 application.

(Please mark 'X' in the Yes/No column as appropriate to indicate your response.)

1. Are research participants classed as **people whose ability to give free and informed consent is in question**? (This may include those under 18 (though see "competent youths"), prisoners, or adults "at risk".) Your attention is drawn to the University's **Safeguarding Code of Practice** and its implications for researchers involving children or adults at risk, including the need for the work to be risk assessed and for researchers to undertake related training. (**Note**: If any of your participants are aged 16 or under, please answer 'Yes' here **and** also answer question 5 below.)	Yes **X**	No
2. By taking part in the research, will participants be at serious risk of criminal prosecution (e.g. by providing information on drug abuse or child abuse)?	Yes	No **X**
3. Does the research involve the **deception** of participants?	Yes	No **X**
4. Does your research raise issues relevant to the **Counter-Terrorism and Security Act (the Prevent duty)**, which seeks to prevent people from being drawn into terrorism? Please see advice on this on our **Best Practice Guidance web page**.	Yes	No **X**

If you have answered 'No' to all of the questions above please go to **Section B**. If you have answered 'Yes' to any question above continue to question 5 below.

5. Is your project covered by a CUREC **approved procedure** (formerly known as "CUREC Protocols")?	Yes **X**	No
If yes, please give research procedure number(s): MSc modus operandi		

If you answered '**Yes'** to **ANY** of questions **1-4**, **and** answered '**No'** to question **5**, **please stop completing this checklist and do not submit it for ethical review**. Instead, please complete the CUREC 2 application form from the CUREC website. Then submit the CUREC 2 form for ethical review.
If you answered 'Yes' to ANY of questions 1-3, **and** answered 'Yes' to question 5, please go on to **Section B**.

SECTION B: Contact details and project description (NB: must be typed not handwritten)

Contact details:

1. **Principal investigator /** supervisor (if student research (title and full name):	
2. Name of student (if student research):	
3. Degree programme, e.g. DPhil, BA, MPhil, BSc, MSc (if student research):	MSc Learning and Teaching
4. Department or Institute name:	Department of Education
5. Address for correspondence (if different from above):	
6. University e-mail (**not** private email) and telephone:	
7. Name and status of others taking part in the project, e.g. third year undergraduate; postdoctoral research assistant:	

Appendix 14: Letter of approval to head teacher

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION



3rd November 2018

Dear *******

I am writing to enquire about conducting research in school this academic year. As you know, I am studying for the Master's in Learning and Teaching at Oxford University, supervised by Gabriel Stylianides. In my final research I will explore how regular low-stakes quizzing can help improve knowledge and retention of year 11 learners for performance in high-stakes examinations.

The research will take place with the year 11 cohort. I am looking to develop ways of giving these learners a structured way of practicing regular recall, and to change the way pupils view testing as a method of revision for high-stakes examinations. A critical element of my research is the collaboration with fellow colleagues to delver an intervention. All teachers in the mathematics department have agreed to collaborate with me on this.

By participating in the research, the school would be contributing to a project that will deepen our understanding of mathematics learning for students with both low and high prior attainment, and so contribute towards developing ways of improving attainment for similar students in the school in the future. It will also contribute to Mathematics education more widely.

I hope to conduct this research between November 2018 and July 2019. As part of the research process, I will be interviewing a sample of pupils, in addition to comparing results from both the November and February mock examination series with the corresponding examination results from the previous year.

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

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

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

I look forward to hearing from you.

Yours sincerely,

Improving the memory and retention of year 11 mathematics learners through regular, low-stakes testing.

University of Oxford, Department of Education

- ☐ We do not wish to participate in this project.
- ☐ We would like to find out more about this project.
- ☐ We would like to take part in this project.

_____

Head teacher's signature

Thank you for your help

References

Agarwal, P., Bain, K., & Chamberlain, P. (2012). The Value of Applied Research: Retrieval Practice Improves Classroom Learning and Recommendations from a Teacher, a Principal, and a Scientist. *Educational Psychology Review, 24*(3), 437-448.

Agarwal, P., Finley, J., Rose, N., & Roediger, H. (2017). Benefits from retrieval practice are greater for students with lower working memory capacity. *Memory, 25*(6), 764-771.

Altrichter, H., Posch, P., & Somekh, B. (1993). *Teachers investigate their work : An introduction to the methods of action research* (Routledge series in investigating schooling). London ; New York: Routledge.

Amlund, J., Kardash, C., & Kulhavy, R. (1986). Repetitive Reading and Recall of Expository Text. *Reading Research Quarterly, 21*(1), 49-58.

Anderson, J., Reder, L., & Simon, H. (1996). Situated Learning and Education1. *Educational Researcher, 25*(4), 5-11.

Annis, L., & Kent Davis, J. (1978). Study Techniques and Cognitive Style: Their Effect on Recall and Recognition. *The Journal of Educational Research, 71*(3), 175-178.

Ashcraft, M. (2002). Math Anxiety: Personal, Educational, and Cognitive Consequences. *Current Directions in Psychological Science, 11*(5), 181-185.

Aubusson, P., Schuck, S., & Burden, K. (2009). Mobile learning for teacher professional learning: benefits, obstacles and issues. *ALT-J*, *17*(3), 233-247.

Baddeley, A., Eysenck, M., & Anderson, M. (2009). *Memory*. Hove, East Sussex ; New York: Psychology Press.

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology, 63*(1), 1-29.

Bandura, A. (1977). Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review, 84*(2), 191-215.

Barnett, S., & Ceci, S. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612-637.

Beidel, D., Turner, S., & Taylor-Ferreira, J. (1999). Teaching Study Skills and Test-Taking Strategies to Elementary School Students: The Testbusters Program. *Behavior Modification, 23*(4), 630-646.

Bell, K., & Limber, J. (2010). Reading Skill, Textbook Marking, and Course Performance. *Literacy Research and Instruction, 49*(1), 56-67.

Bernacki, Byrnes, & Cromley. (2012). The effects of achievement goals and self-regulated learning behaviors on reading comprehension in technology-enhanced learning environments. *Contemporary Educational Psychology, 37*(2), 148-161.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. *From learning processes to cognitive processes: Essays in honor of William K. Estes*, *2*, 35-67.

Bjork, R.A. (1994). Memory and metamemory considerations in the training of human beings. InMetcalfe, J. & Shimamura, A. (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Bjork, R. A. (1999). F 5 assessing our own competence: Heuristics and illusions.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*, *2*(59-68).

Bjork, R., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology,64*(1), 417-444.

Black, P., & Wiliam, D. (1998). *Inside the black box : Raising standards through classroom assessment*. London: King's College London School of Education.

Bol, L., Warkentin, R. W., Nunnery, J. A., & O'Connell, A. A. (1999). College students' study activities and their relationship to study context, reference course, and achievement. *College Student Journal*, 33, 608-622.

Bråten, I., & Samuelstuen, M. (2007). Measuring strategic processing: Comparing task-specific self-reports to traces. *Metacognition and Learning, 2*(1), 1-20.

Bull, R, & Scerif, Gaia. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*, 273-293.

Butler, A. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36*(5), 1118-1133.

Callender, & Mcdaniel. (2009). The limited benefits of rereading educational texts. *Contemporary Educational Psychology, 34*(1), 30-41.

Censabella, S., & Noël, M. (2008). The Inhibition Capacities of Children with Mathematical Disabilities. *Child Neuropsychology, 14*(1), 1-20.

Changes to GCSE's and A-Levels are damaging students mental health and increasing teachers' workload. (2018). Retrieved 16 April 2019, from National Union of Teachers website: https://neu.org.uk/assessment/changes-gcses-and-levels-are-damaging-students-mental-health-and-increasing-teachers

Chartrand, T., & Bargh, J. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893-910.

Christodoulou, D. (2013). *Seven myths about education* (Ebook central). Oxfordshire, England ; New York.

Clark, C. M. (1992). The restructuring of meaning: An analysis of the impact of context on transformational learning. 33rd Annual Adult Education Research Conference Proceedings (pp. 31-36). University of Saskatchewan, Saskatoon.

Cohen, D. (2013) Revision techniques – the good, the OK and the useless, BBC news 18 May 2013. Available from < https://www.bbc.co.uk/news/health-22565912> [Accessed 24 August 2019]

Cohen, L., & Manion, L. (1980). *Research methods in education*. London: Croom Helm.

Congleton, A., & Rajaram, S. (2011). The influence of learning history on collaborative and individual retrieval. *J. Exp. Psych: Gen*.

Coser, L., & Hirsch, E. (1987). Cultural Literacy: What Every American Needs to Know. *Science,236*(4804), 973.

Dadds, M., Hart, S., & Crotty, T. (2001). *Doing practitioner research differently*. London: RoutledgeFalmer.

Das, S. (2019) Damien Hinds: Exam stress is part of life, The Sunday Times, 28 April 2019. Available from < https://www.thetimes.co.uk/article/damian-hinds-exam-stress-is-part-of-life-jsrvpklcw#> [Accessed 20 July 2019]

Delamont, S. (2002) Fieldwork in Educational Settings: Methods, pitfalls and perspectives. London: Routledge.

Dembo, M., & Seli, H. (2004). Students' Resistance to Change in Learning Strategies Courses. *Journal of Developmental Education, 27*(3), 2.

Denscombe, M. (2017). *The good research guide : For small-scale social research projects* (Sixth ed., Open UP study skills). London.

Didau, D. (2013) Testing and Assessment – Have we been doing the right things fo the wrong reasons? 16 June 2013. Available from https://learningspy.co.uk/assessment/afl-have-we-been-doing-the-right-things-for-the-wrong-reasons-2/ [Accessed 8 February 2019]

Dubinsky, J., Roehrig, G., & Varma, S. (2013). Infusing Neuroscience Into Teacher Professional Development. *Educational Researcher, 42*(6), 317-329.

Duckworth, A., Peterson, C., Matthews, M., & Kelly, D. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087-101.

Dunlosky, J., Rawson, K., Marsh, E., Nathan, M., & Willingham, D. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest, 14*(1), 4-58.

Ebbinghaus H (1885) Über das Gedächtnis. Leipzig: Dunker.

Education Endowment Foundation Teacher Toolkit, (2013), *Sutton Trust EEF Teaching and Learning Toolkit.* Retrieved from: https://www.cheshireeast.gov.uk/PDF/Sutton_Trust_%20EEF_T_&_L_Toolkit_July_13.pdf

Education Endowment Foundation Teacher Toolkit, (2018), *Sutton Trust EEF Teaching and Learning Toolkit.* Retrieved from: https://educationendowmentfoundation.org.uk/public/files/Presentations/Publications/Metacognition/EEF_Metacognition_and_self-regulated_learning.pdf

Einstein, G., Mullet, H., & Harrison, T. (2012). The Testing Effect: Illustrating a Fundamental Concept and Changing Study Strategies. *Teaching of Psychology, 39*(3), 190-193.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of learning and motivation*, *44*, 145-200

Entwistle, N., & McCune, V. (2004). The Conceptual Bases of Study Strategy Inventories. *Educational Psychology Review, 16*(4), 325-345.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist,34*(10), 906-911.

Foddy, W.H. (1994) Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research. New Edition, Cambridge University Press, Cambridge, UK.

Gates, A. (1917). *Recitation as a factor in memorizing*. University printing office, Columbia University.

Geary, Hoard, Byrd-Craven, & Catherine Desoto. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology, 88*(2), 121-151.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing:

Failure in the self-assessment of comprehension. *Memory and Cognition*, 10, 597-602.

Gurung, R. (2005). How do students really study (and does it matter)? *Teaching Of Psychology, 32*(4), 239-241.

Hargreaves, A. (2000). Four Ages of Professionalism and Professional Learning. *Teachers and Teaching, 6*(2), 151-182.

Hartwig, M., & Dunlosky, K. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review, 19*(1), 126-134.

Hattie, J A C (2011) – Visible Learning for Teachers : Maximising Impact on Learning (London: Routledge)

Heritage, M. (2018) Making Assessment Work for Teachers. Educational Measurement: Issues and Practice, 37 (1), 39–41.

Hirsch, E.D., Jr. The Knowledge Deficit: Closing the Shocking Education Gap for American Children. Houghton. Apr. 2006. 192p.

Hofer, B., Yu, S., & Pintrich, P. R. (1998). Teaching college students to be self-regulated learners. In D. H. Schunk & B. J. Zimmerman (Eds.), Self-regulated learning: From teaching to self-reflective practice (pp. 57–85). New York: Guilford.

Ireson, J., Hallam, S., & Plewis, I. (2001). Ability grouping in secondary schools: Effects on pupils' self-concepts. *The British Journal of Educational Psychology, 71*(Pt 2), 315-26.

Jaap M J Murre, & Joeri Dros. (2015). Replication and Analysis of Ebbinghaus' Forgetting Curve. *PLoS ONE, 10*(7), E0120644.

Paul A. Howard-Jones. (2014). Neuroscience and education: Myths and messages. *Nature Reviews Neuroscience, 15*(12), 817-81724.

Izawa, C. (1966). Reinforcement-test sequences in paired-associate learning. *Psychological Reports*, *18*(3), 879-919.

Jacoby, L. L., Bjork, R. A., & Kelley, C. M. (1994). Illusions of comprehension, competence, and remembering. In D. Druckman & R. A. Bjork (Eds.), Learning, remembering, believing: Enhancing human performance (pp. 57–80). Washington, DC: National Academy Press.

Johnson, R., Onwuegbuzie, A., & Turner, L. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research, 1*(2), 112-133.

Karpicke, J., & Roediger, H. (2007). Expanding Retrieval Practice Promotes Short-Term Retention, but Equally Spaced Retrieval Enhances Long-Term Retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(4), 704-719.

Karpicke, Jeffrey D., & Roediger, Henry L. III. (2008). Critical Importance of Retrieval for Learning. *Science, 319*(5865), 966-968.

Karpicke, J., Butler, A., & Roediger Iii, H. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory, 17*(4), 471-479.

Karpicke, J., & Aue, D. (2015). The Testing Effect Is Alive and Well with Complex Materials. *Educational Psychology Review, 27*(2), 317-326.

King, A. (1992). Facilitating Elaborative Learning Through Guided Student-Generated Questioning. *Educational Psychologist, 27*(1), 111-126.

Kornell, N., & Bjork, R. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review, 14*(2), 219-224.

Kornell, N., & Son, L. (2009). Learners' choices and beliefs about self-testing. *Memory, 17*(5), 493-501.

LeFevre, J.-A., DeStefano, D., Coleman, B., & Shanahan, T. (2005). Mathematical cognition and working memory. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 361-377). New York, NY, US: Psychology Press.

Lepper, M. (1988). Motivational Considerations in the Study of Instruction. *Cognition and Instruction, 5*(4), 289-309.

Lonka, Lindblom-Ylänne, & Maury. (1994). The effect of study strategies on learning from text. *Learning and Instruction, 4*(3), 253-271.

Masson, & Mcdaniel. (1981). The role of organizational processes in long-term retention. *Journal of Experimental Psychology: Human Learning and Memory, 7*(2), 100-110.

Matthews, G. S., Grier, R. E., Warm, J. A., Campbell, S., Falconer, S., Joyner, L., . . . Gilliland, K. (2002). Fundamental Dimensions of Subjective State in Performance Settings : Task Engagement, Distress, and Worry. *Emotion, 2*(4), 315-340.

Mayer, R. (1996). Learning strategies for making sense out of expository text: The SOI model for guiding three cognitive processes in knowledge construction. *Educational Psychology Review, 8*(4), 357-371.

Mcandrew, M., Morrow, C., Atiyeh, L., & Pierre, G. (2016). Dental Student Study Strategies: Are Self-Testing and Scheduling Related to Academic Performance? *Journal Of Dental Education, 80*(5), 542-552.

Mccabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition, 39*(3), 462-476.

Mcdaniel, M., Agarwal, P., Huelser, B., Mcdermott, K., & Roediger, H. (2011). Test-Enhanced Learning in a Middle School Science Classroom: The Effects of Quiz Frequency and Placement. *Journal of Educational Psychology, 103*(2), 399-414.

McDaniel, M., Thomas, R., Agarwal, P., McDermott, K., & Roediger, H. (2013). Quizzing in Middle- School Science: Successful Transfer Performance on Classroom

McGeoch, J. (1932). Forgetting and the law of disuse. *Psychological Review, 39*(4), 352-370.

Meijer, J., & Riemersma, F. (2002). Teaching and testing mathematical problem solving by offering optional assistance. *Instructional Science, 30*(3), 187-220.

Miller, G. A., Eugene, G., & Pribram, K. H. (1960). Plans and the Structure of Behaviour. In *Systems Research for Behavioral Science* (pp. 369-382). Routledge.

Nelson, T., & Leonesio, R. (1988). Allocation of self-paced study time and the "labor-in-vain effect". *Journal of Experimental Psychology. Learning, Memory, and Cognition, 14*(4), 676-86.

Newton, L. (1990). Overconfidence in the communication of intent: Heard and unheard melodies. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Nist, S., & Kirby, K. (1989). THE TEXT MARKING PATTERNS OF COLLEGE STUDENTS. *Reading Psychology, 10*(4), 321-338.

Oakes, J. (1982). The reproduction of inequity: The content of secondary school tracking. *The Urban Review, 14*(2), 107-120.

Oakley, B. (2015). Learning How We Learn: In an unlikely confluence, bioengineers are finding fertile ground in research and teaching about learning. *IEEE Pulse, 6*(2), 37-41.

Ofsted (2019). Common Inspection framework. Available from: https://www.gov.uk/government/publications/education-inspection-framework [Accessed 9 June 2019].

Onwuegbuzie, A., Leech, N., & Collins, K. (2012). Qualitative Analysis Techniques for the Review of the Literature. *Qualitative Report, 17*, Qualitative Report, 2012, Vol.17.

Panadero, E., Klug, J., & Järvelä, S. (2016). Third wave of measurement in the self-regulated learning field ; when measurement and intervention come hand in hand. *Scandinavian Journal of Educational Research, 60*, 723-735.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning styles concepts and evidence. *Psychological Science in the Public Interest, Supplement, 9*(3), 105-119.

Pintrich, P., Smith, D., Garcia, T., & Mckeachie, W. (1993). Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (Mslq). *Educational and Psychological Measurement, 53*(3), 801-813.

Pring, R. (2000). The 'False Dualism' of Educational Research. *Journal of Philosophy of Education, 34*(2), 247-260.

Purdie, N., & Hattie, J. (1999). The Relationship between Study Skills and Learning Outcomes: A Meta-Analysis. *Australian Journal of Education, 43*(1), 72-86.

Pyc, M., & Rawson, K. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science (New York, N.Y.), 330*(6002), 335.

Rickards, J., & Denner, P. (1979). Depressive effects of underlining and adjunct questions on children's recall of text. *Instructional Science, 8*(1), 81-90.

Rinehart, S. (1986). Some Effects of Summarization Training on Reading and Studying. *Reading Research Quarterly, 21*(4), 422-438.

Roediger, H. (1985). Remembering Ebbinghaus. *Contemporary Psychology: A Journal of Reviews, 30*(7), 519-523.

Roediger, & Butler. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20-27.

Roediger III, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. In *Psychology of learning and motivation* (Vol. 55, pp. 1-36). Academic Press.

Rovers, Sanne F. E., Stalmeijer, Renee E., Van Merrienboer, Jeroen J. G., Savelberg, Hans H. C. M., De Bruin, Anique B. H., Promovendi SHE, . . . Nutrition Movement Sciences. (2018). How and Why Do Students Use Learning Strategies? A Mixed Methods Study on Learning Strategies and Desirable Difficulties With Effective Strategy Users. *Frontiers in Psychology, 9*, Urn:issn:1664-1078.

Samson, G. (1985). Effects of Training in Test-Taking Skills on Achievement Test Performance: A Quantitative Synthesis. *The Journal of Educational Research, 78*(5), 261-266.

Schoenfeld, A. (1987). *Cognitive science and mathematics education*. Hillsdale, N.J. ; London: Erlbaum.

Schunk, D. H., & Zimmerman, B. J. (Eds.). (1994). Self-regulation of learning and performance: Issues and educational applications. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Shaughnessy, J. J., & Zechmeister, E. B. (1992). Memory-monitoring accuracy as influenced by the distribution of retrieval practice. Bulletin of the Psychonomic Society, 30, 125–128.

Skemp, R. (1976). Relational Understanding and Instrumental Understanding. *Mathematics Teaching,*Mathematics Teaching, 1976.

Skinner, B. (1985). Cognitive science and behaviourism. *British Journal of Psychology, 76*(3), 291-301.

Soderstrom, N., & Bjork, R. (2015). Learning Versus Performance: An Integrative Review. *Perspectives on Psychological Science, 10*(2), 176-199.

Stallings, J. A., & Kaskowitz, D. H. (1974). Follow Through Classroom Observation Evaluation 1972-1973. SRI Project URU-7370.

Stordahl, K., & Christensen, C. (1956). The Effect of Study Techniques on Comprehension and Retention. *The Journal of Educational Research, 49*(8), 561-570.

Sundre, & Kitsantas. (2004). An exploration of the psychology of the examinee: Can examinee self-regulation and test-taking motivation predict consequential and non-consequential test performance? *Contemporary Educational Psychology, 29*(1), 6-26.

Swanson, H., & Jerman, O. (2006). Math Disabilities: A Selective Meta-Analysis of the Literature. *Review of Educational Research, 76*(2), 249-274.

Tait, H., & Entwistle, N. (1996). Identifying students at risk through ineffective study strategies. *Higher Education, 31*(1), 97-116.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research : Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks ; London: Sage.

Thorndike, E., & Woodworth, R. (1901). The Influence of Improvement in One Mental Function upon the Efficiency of Other Functions. *Psychological Review, 8*, 247.

Todd, W., & Kessler, C. (1971). Influence of Response Mode, Sex, Reading Ability, and Level of Difficulty on Four Measures of Recall of Meaningful Written Material. *Journal of Educational Psychology,62*(3), 229-234.

Trkulja, J. (2013) GCSE overhaul in England made final by Ofqual. BBC news website 1 November 2013. Available from: <https://www.bbc.co.uk/news/education-24759476> [Accessed 16 May 2019]

Van Gog, T., & Sweller, J. (2015). Not New, but Nearly Forgotten: The Testing Effect Decreases or even Disappears as the Complexity of Learning Materials Increases. *Educational Psychology Review, 27*(2), 247-264.

Wegner, D., & Wheatley, T. (1999). Apparent mental causation. Sources of the experience of will. *The American Psychologist, 54*(7), 480-492.

Wheeler, M., & Roediger, H. (1992). Disparate Effects of Repeated Testing: Reconciling Ballard's (1913) and Bartlett's (1932) Results. *Psychological Science, 3*(4), 240-245.

Wheeler, M., Ewers, M., & Buonanno, J. (2003). Different rates of forgetting following study versus test trials. *Memory, 11*(6), 571-580.

Willingham, D. (2002). How We Learn. Ask the Cognitive Scientist: Allocating Student Study Time. "Massed versus "Distributed" Practice. American Educator. 26.

Willingham, D. (2009). *Why don't students like school? : A cognitive scientist answers questions about how the mind works and what it means for the classroom* (1st ed.). San Francisco, CA: Jossey-Bass.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *The educational psychology series. Metacognition in educational theory and practice* (pp. 277-304). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Wooldridge, Bugg, Mcdaniel, & Liu. (2014). The testing effect with authentic educational materials: A cautionary note. *Journal of Applied Research in Memory and Cognition, 3*(3), 214-221.

Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), Self-efficacy in changing societies (pp. 202–231). New York: Cambridge Univ. Press

Zimmerman, B. (1998). Academic studing and the development of personal skill: A self-regulatory perspective. *Educational Psychologist,33*(2-3), 73-86.

Zimmerman, Barry J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory into Practice, 41*(2), 64-72.