Faculty of Cognitive Sciences and Human Development

The Effects of Retrieval Practice Across Levels of Thinking and
Retention Interval on Reading Comprehension

Corrin Alicia anak Nero

Master of Science
2022

The Effects of Retrieval Practice Across Levels of Thinking and Retention
Interval on Reading Comprehension

Corrin Alicia anak Nero

A thesis submitted

In fulfillment of the requirements for the degree of Master of Science

(Cognitive Science)

Faculty of Cognitive Sciences and Human Development
UNIVERSITI MALAYSIA SARAWAK
2022

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

………………………………

Signature

Name: Corrin Alicia anak Nero

Matric No.:   19020133

Faculty of Cognitive Sciences and Human Development

Universiti Malaysia Sarawak

Date : 15 September 2022

ACKNOWLEDGEMENT

Before anything else, all praise and thanks to God for blessing my journey in completing this work. I thank Him for giving me the strength during difficult times and wisdom, when I lacked it.

I would like to express my sincere appreciation to my supervisor, Dr Norehan Zulkiply for her time, effort, and guidance in the completion of this research. I thank her for the words of encouragement and for the invaluable advice she had given throughout this journey, in both academic research as well as life.

Special thanks to the two most important figures in my life—my beloved parents, Mr Nero Madi and Mdm Bibi Chia for their unwavering love and support. Needless to say, I would not be where I am today without them. My gratitude also extends to my sister, Daine Esther, for her continuous support.

My appreciation also goes out to my friends who have been there for me in my ups and downs. Thank you for continuously motivating, advising, and tirelessly listening to my rants (you know who you are).

To the dramas, artistes, and fandoms that I have got myself into, thank you for being a constant space of comfort especially during stressful times.

Finally, I would like to thank all the students for participating in the study. The completion of this research would not have been possible without their participation.

ABSTRACT

The present study examined the effect of different types of retrieval practice on reading comprehension across levels of thinking and retention interval in students. One hundred undergraduates divided into two retention interval groups (short and long retention interval) were asked to read a passage on a topic in Cognitive Psychology and were then required to engage in two types of retrieval practice learning strategy (production test and recognition test) and different levels of thinking (lower-order thinking and higher-order thinking). A three-way mixed ANOVA statistical test was used to analyse the data and found no significant difference in reading comprehension across the different types of retrieval practice, suggesting that the performance when using the recognition test was equivalent to when using the production test. The difference in reading comprehension between the different types of retention interval also was not observed, indicating that students in the short retention interval group retained just as much information as those in the long retention interval group. Additionally, the present study observed a significant difference in students' reading comprehension between different levels of thinking, signifying that the students' performance for the lower-order thinking questions was better as compared to the higher-order thinking questions. The present finding contributed to the existing body of knowledge in which it suggested that the performance in reading comprehension when using a recognition test, particularly a well-constructed one, with competitive alternatives was equivalent to when using a production test.

Keywords: Retrieval practice, levels of thinking, retention interval, reading comprehension, undergraduates

Kesan Latihan Perolehan Semula, Tahap Pemikiran dan Selang Pengekalan terhadap Pemahaman Membaca

ABSTRAK

Kajian ini mengkaji kesan latihan perolehan semula, tahap pemikiran dan selang pengekalan terhadap pemahaman membaca dalam kalangan pelajar prasiswazah. Seratus peserta dibahagikan kepada dua kumpulan (selang pengekalan pendek dan selang pengekalan panjang) dan mereka diminta untuk membaca petikan tentang satu topik dalam subjek Psikologi Kognitif. Peserta kajian kemudiannya dikehendaki untuk mengulangkaji petikan dengan menggunakan dua jenis strategi pembelajaran latihan perolehan semula (ujian penjanaan jawapan secara bebas dan ujian objektif aneka pilihan) dan melibatkan tahap pemikiran yang berbeza (pemikiran aras rendah dan pemikiran aras tinggi). Ujian statistik ANOVA campuran tiga hala digunakan untuk menganalisis data dan didapati tiada perbezaan signifikan dalam pemahaman membaca antara pelbagai jenis latihan perolehan semula, sekaligus menunjukkan bahawa prestasi pemahaman membaca pelajar adalah sama apabila menggunakan ujian objektif aneka pilihan dan ujian penjanaan jawapan secara bebas. Hasil analisis ujian statistik juga tidak menemui perbezaan signifikan dalam pemahaman membaca pelajar antara pelbagai jenis selang pengekalan sekaligus menunjukkan bahawa jumlah maklumat yang dikekalkan oleh pelajar adalah sama untuk selang pengekalan pendek dan panjang. Di samping itu, kajian ini menemui perbezaan yang signifikan dalam pemahaman membaca pelajar antara tahap pemikiran yang berbeza, menandakan bahawa prestasi pemahaman pelajar untuk soalan pemikiran aras rendah adalah lebih baik berbanding dengan soalan pemikiran aras tinggi. Penemuan dalam kajian ini menyumbang kepada literatur sedia ada di mana dapatan kajian ini mencadangkan bahawa prestasi dalam pemahaman membaca apabila menggunakan ujian objektif aneka

pilihan, terutamanya jika soalannya dibina dengan baik, dengan alternatif yang kompetitif adalah sama apabila menggunakan ujian penjanaan jawapan secara bebas.

Kata kunci: *Latihan perolehan semula, tahap pemikiran, selang pengekalan, pemahaman membaca, prasiswazah*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

HOT   Higher-Order Thinking

LOT   Lower-Order Thinking

MCQ   Multiple-Choice Question

MCQs   Multiple-Choice Questions

CHAPTER 1

INTRODUCTION

1.1   Introduction

Chapter One presents the general picture of this research. It describes the background of the study, highlights the problem statement, and defines the research objectives as well as research questions. This chapter also includes the definition of terms used, conceptual framework and the significance of the study.

1.2   Background of Study

Tests in educational setting are conventionally used as a tool in the assessment of learning (Amin & Malik, 2013; Brame & Biel, 2015; Moreira et al., 2019) as well as to provide a summarised picture of an individual's knowledge (Stenlund et al., 2016). However, a growing number of studies have revealed that tests can do more than just measuring one's knowledge, that is, it can boost learning as well as promotes long-term memory retention (Stenlund et al., 2016; Karpicke, 2017; Moreira et al., 2019). Studies that examined the effects of test usually compares the condition where learners take test or engage in retrieval practice task on previously studied materials to the condition where learners restudy or reread the previously studied materials. Experimental procedures in examining this effect usually involved the following phases: study phase, retrieval practice task phase and final test phase. Participants were firstly required to study a set of material, for example word pairs, word lists or reading passages and then instructed to take part in a retrieval practice task, either in the form of a free recall test, a multiple-choice question (MCQ) or a short-answer test. After engaging in the retrieval practice task, they have to complete a final test, either immediately or after a delay (e.g., hours, days, or weeks) to measure their performance

in terms of their understanding of the studied material or how much information they can recall from their memory.

Several studies have shown that when learners engage in the practice of retrieving previously studied information, it can result in greater long-term memory retention as compared to merely rereading or restudying that particular information (Roediger & Karpicke, 2006; Karpicke & Roediger, 2008; Stenlund et al., 2016; Koh et al., 2018). In both experiments conducted by Roediger and Karpicke (2006), students were required to study reading passages and then undertake one or three immediate free recall tests without feedback or restudy the material the same number of times as the students who took the tests. Students then took a final retention test five minutes, two days, or one week later. Restudying was included as one of the conditions based on the idea that the students were represented with the information once again. The results, particularly the second experiment which examined repeated testing versus repeated studying revealed that when the final test was given after five minutes, restudying improved recall compared to repeated testing. However, on the delayed tests, testing was revealed to have produced greater memory retention than studying, although repeated studying increased students' confidence in their ability to recall the learned material. Similarly, Stenlund et al.'s (2016) study which compared the effects of repeated practice testing with short-answer items on short- and long-term memory retention to repeated practice testing with multiple-choice questions, as well as rereading the material found that both test formats (short-answer and MCQ) enhanced short- and long-term memory relative to rereading. Rereading or restudying is usually used as the comparison due to the basic assumption that learning happens when learners study and encode material and based on such an assumption, rereading or restudying should increase learning (Karpicke & Roediger, 2008). However, such is not always the case as several studies revealed that using

test or retrieval practice is more effective for memory retention as it involves effortful retrieval, whereas studying or reading the material simply involves encoding (Roediger & Karpicke, 2006; Karpicke & Roediger, 2008).

When one studies and then proceeds to retrieve the previously learned information, it enhances learning and promotes long-term memory retention more than merely restudying the material. This phenomenon is known as the testing effect (McDaniel, Roediger, et al., 2007; Stenlund et al., 2016; Moreira et al., 2019). This phenomenon is also commonly referred to as the retrieval practice effect (Karpicke & Roediger, 2007a; Karpicke, 2017) or retrieval-based learning (Karpicke, 2017; Bae et al., 2019). The explanation behind the effectiveness of this effect is based on several theories. One of the theories is the retrieval effort hypothesis which states that a more difficult and more effortful processing has greater beneficial effects on memory (Gardiner et al., 1973; Pyc & Rawson, 2009). The act of taking a test requires the use of effortful retrieval in which learners have to actively retrieve information from memory (Karpicke & Roediger, 2008). When information is successfully retrieved from memory, the representation in memory is changed in a way that it becomes more likely to be recalled in the future (Bjork, 1975) thus, yielding greater benefit compared to repeatedly studying the materials (Roediger & Karpicke, 2006). Another theory that is commonly used in explaining the retrieval practice effect is transfer-appropriate processing. This theory states that when the cognitive process that a student engages in during the initial learning activity is identical with the processing required during the final assessment, it leads to a better performance in learning and memory tasks (Morris et al., 1977; Kolers & Roediger, 1984). This is based on the assumption that the positive results on long-term memory retention are due to the resemblance of cognitive processes during learning tasks and transfer tasks (Endres & Renkl, 2015), which further explains why test or retrieval

practice has an advantage over restudying or rereading. This advantage exists due to the cognitive processes that are present during the transfer tasks being more similar to the ones used during initial testing, rather than restudying or rereading (e.g., Thomas & McDaniel, 2007). Another theory to explain the retrieval practice effect is the elaborative-retrieval account. This theory explains that semantic elaboration takes place during the process of retrieval which will improve subsequent retrieval (Karpicke, 2017). The idea behind this theory is that when one retrieves a target from a given cue, it activates additional information that is semantically associated to the cue and this information is incorporated with the successfully recalled target to create an elaborate memory trace that becomes easily remembered at a later time (Carpenter & Yeung, 2017; Karpicke, 2017). When one restudies or rereads material on the other hand, the production of additional mediators does not necessarily happen because there is no activity of retrieving the target word during restudy trials (Carpenter, 2009) hence explaining why retrieval practice results in better performance than restudying. Across a wide number of past studies on the retrieval practice effect, consistent findings have reported that retrieval practice has greater beneficial effects than rereading or restudying (Carpenter et al., 2006; McDaniel, Roediger, et al., 2007; Rowland, 2014; Karpicke, 2017) thus indicating that it is an established fact that retrieval practice enhances memory retention relative to rereading or restudying.

It was noted in Greving and Richter's (2018) study that although there is convincing evidence pertaining to the beneficial effects of using retrieval practice, the circumstances in which the effects occur are still open for discussion. One such example is the generalisability of the results, especially in classroom contexts. Therefore, considering that there have been consistent findings on various contexts of studies which reported that retrieval practice had greater benefits than rereading, the present study aims to focus on the effects of the two

different types of retrieval practice format. There are two commonly used test or retrieval practice formats in examining testing effect which are the recognition test and the production test (Butler & Roediger, 2007; Pyc & Rawson, 2009; Larsen & Butler, 2013; Greving & Richter, 2018). The production test is usually in the form of free recall, short-answer essay, and fill-in-the-blanks, whereas the recognition test is in the form of multiple-choice questions (MCQs) and "true or false" (Larsen & Butler, 2013; Moreira et al., 2019). Moreira et al. (2019) noted that the differences with regard to the effectiveness of the different types of tests in drawing out the testing effect remains a question for both laboratory and classroom contexts, wherein there are mixed findings with regard to the effectiveness of different types of test format in producing the testing effect. In other words, some studies reported that tests which involve production task would provide greater benefit than recognition task (Butler & Roediger, 2007; Greving & Richter, 2018) whereas some studies reported otherwise (Little et al., 2012; Smith & Karpicke, 2014; Adesope et al., 2017). In Greving and Richter's (2018) study which examined the testing effect in a classroom setting whereby the participants were asked to either answer short-answer questions, MCQs or read summarising statements about lecture content, it was found that short-answer testing may have provided an additional advantage in learning in higher education contexts compared to multiple-choice testing. The reason to such finding can be explained in terms of cognitive effort whereby questions that prompted effortful retrieval were likely to produce greater retrieval practice effects, of which in this case would be the short-answer test. However, there are some studies which revealed that MCQs are also as effective as production test in eliciting effects. Smith and Karpicke's (2014) study found that there was little or no advantages in answering short-answer questions compared to MCQs. Such finding appeared due to the unsuccessful retrievals when using short-answer items whereby most of the time participants seemed to perform better in MCQs

than in short-answer tests. They explained that even though a more difficult test is said to be more advantageous, success of retrieval is equally as important for subsequent performance (Smith & Karpicke, 2014).

In terms of retrieval practice and levels of thinking, there has been an increasing number of studies which examine the effects of retrieval practice on different levels of thinking, that is, lower-order thinking and higher-order thinking (Butler, 2010; McDaniel et al., 2013; Carpenter et al., 2016; Dobson et al., 2018; Agarwal, 2019). In particular, these studies aim to determine whether transfer of knowledge from one level to another level (e.g., lower-level thinking to higher-level thinking or vice versa) would occur if students were instructed to engage in a learning task using the retrieval practice strategy. For example, participants were required to study a set of text passages and then were assigned to the experimental manipulation conditions where they either restudied the passage or engaged in a retrieval practice strategy (i.e., lower-order thinking test or higher-order thinking test). After a delay, they were given a final test that comprised both lower-order and higher-order thinking questions to observe the transfer of knowledge from one knowledge domain to another and to see whether the performance on the lower-order thinking task would enhance performance on the higher-order thinking task. Findings from the studies which examined this matter showed that even though the levels of thinking are usually said to follow a hierarchy, superior performance on lower-order thinking does not necessarily enhance performance of higher-order thinking tasks (e.g., McDaniel et al., 2013, Dobson et al., 2018, Agarwal, 2019). However, considering the majority of previous studies on retrieval practice and levels of thinking, there has yet to be a study that mainly investigates the effect of different types of question format used in retrieval practice (production test versus recognition test) on different levels of thinking. For instance, most studies that observe the

effect of retrieval practice on levels of thinking usually examine the effect of the transfer of thinking skills and most of the time they used only one question format, which is either a production test or a recognition test (Butler, 2010; Dobson et al., 2018; Agarwal, 2019). It is important to study the effects of both types of question formats especially in terms of addressing the effect in an educational setting since different question formats, particularly short-answer questions as well as MCQs are usually implemented in classrooms. Also, there are still ongoing arguments involving the issue of how MCQs and short-answer questions differ in terms of the success rate in the retrieval process and particularly its effect on learning (Karpicke, 2017). Therefore, this underlines the need for further investigation to find out whether there is any effect of the different types of question format on levels of thinking among students.

A growing number of studies investigated the retrieval practice effect and examined the effect on different lengths of delay have shown that when one adopts retrieval practice as a learning strategy, it benefits memory retention (Roediger & Karpicke, 2006; Wiklund-Hörnqvist et al., 2014; Karpicke, 2017). However, in terms of observing the retrieval practice effect especially involving the different types of retrieval practice formats, most studies did not observe the testing benefits across different retention intervals, that is, they either observed the effects of different types of retrieval practice format on short retention intervals or on long retention intervals (e.g., Butler & Roediger, 2007; Little et al., 2012; Smith & Karpicke, 2014). Therefore, not much is known about the long-term retrieval practice effect on the retention of production task and recognition task items (Stenlund et al., 2016). One of the very few studies that observed a retrieval practice effect using different types of formats on different retention intervals was conducted by Stenlund et al. (2016) who revealed that retrieval practice using short-answer items is more beneficial for long-term retention. On the

contrary, a study by Little et al. (2012) found that a well-constructed recognition test with competitive alternatives is also capable of drawing out productive retrieval processes. However, it observed the effect using short retention intervals, therefore, it remains unclear whether the benefits of multiple-choice testing can be retained for longer periods of time.

1.3   Problem Statement

Across a wide number of studies on retrieval practice effect that were conducted over the years, it is an established fact that retrieval practice enhances retention compared to rereading or restudying (Carpenter et al., 2006; Roediger & Karpicke, 2006; McDaniel, Roediger, et al., 2007; Rowland, 2014; Karpicke, 2017). When information is successfully retrieved from memory, the representation in memory is changed in a way that it becomes highly likely to be recalled more in the future (Bjork, 1975) thus, yielding a greater benefit compared to repeatedly studying or reading the materials which only involves encoding (Roediger & Karpicke, 2006). Previous studies stated that the mental process involved in actively retrieving information from memory when one is taking a test, is the mechanism driving the retrieval practice effect, wherein when one actively retrieves information, it strengthens retrieval routes in memory, hence permitting easier retrieval of the particular information in the future. (McDaniel & Masson, 1985; Carrier & Pashler, 1992; Carpenter & DeLosh, 2006; McDaniel, Roediger, et al., 2007; McConnell et al., 2018). Although restudying or rereading the passage exposes the students to the entire set of information again, engaging in retrieval practice allows the practice of the skills necessary for upcoming tests, resulting in greater performance after a delay (Roediger & Karpicke, 2006). While there is compelling evidence for the effects of retrieval practice, there are still questions with regard to the circumstances under which the effects arise, particularly in generalising the benefits of this effect in classroom setting (Greving & Richter, 2018). Therefore, in

consideration of the fact that there have been consistent findings on various contexts of studies which reported that retrieval practice has greater benefits than repeated reading or studying, the present study intends to focus mainly on the effects of the two types of retrieval practice format.

As mentioned earlier, few studies that looked into the different types of test (production test versus recognition test) have produced mixed findings. Some studies have found that tests which involve a production task will provide greater benefit compared to a recognition task (Butler & Roediger, 2007; Pyc & Rawson, 2009; Greving & Richter, 2018) whereas others have not (Little et al., 2012; Smith & Karpicke, 2014). A study by Greving and Richter (2018) which observed the retrieval practice effect in a classroom setting whereby the participants were required to either answer short-answer questions, MCQs or read summarising statements about lecture content discovered that the short-answer test may have additional benefits in learning compared to the multiple-choice test. Such a finding can be attributed to the role of cognitive effort by which questions that induce effortful retrieval are likely to produce greater retrieval practice effects, and in this case these would be the short-answer tests. On the contrary, some studies showed that MCQs are also effective in drawing out the retrieval practice effect. For example, Smith and Karpicke's (2014) study found that there was little or no advantages in answering short-answer tests over MCQs. This finding was based on the unsuccessful retrievals when using short-answer questions. The researchers explained that even though a more difficult test format is said to be more beneficial, the retrieval success rate is equally as crucial for later performance (Smith & Karpicke, 2014).

There has been a growing number of studies of the retrieval practice effect on different levels of thinking, which specifically examine the transfer of knowledge from lower-order thinking to higher-order thinking or vice versa (e.g., McDaniel et al., 2013, Dobson et al., 2018, Agarwal, 2019). However, looking at these studies, there has yet to be a study that mainly investigates the effect of different types of question format used in retrieval practice (production test versus recognition test) on different levels of thinking. For example, most studies that examined the retrieval practice effect on levels of thinking usually investigated the effect in terms of the transfer of thinking skills from one level to another and they usually used only one question format, which was either production test or a recognition test (e.g., Butler, 2010; Dobson et al., 2018; Agarwal, 2019). Considering the fact that different question formats, notably short-answer questions and MCQs, are commonly adopted in classrooms, it is important to explore the effect of both types of question formats, particularly in terms of addressing the retrieval practice effect in educational contexts. Besides that, student's higher-order thinking skills development is considered as an important element of education and recently, educators as well as scientists have been actively developing strategies and instructional approaches that aim to increase higher-order thinking (Agarwal, 2019). Thus, further examining the effects of different question formats on levels of thinking, can shed light on how educators can design learning materials that will improve students' performance and learning experience either for lower-order thinking questions or higher-order thinking questions.

The existing body of literature that observes the effect of retrieval practice on different lengths of delay has shown its benefits for memory retention which further proves that there is an added advantage when one uses retrieval practice as a learning strategy over rereading or restudying (Roediger & Karpicke, 2006; Wiklund-Hörnqvist et al., 2014;