University of Reading

THE EFFECT OF MODE OF PRESENTATION, COGNITIVE LOAD, AND INDIVIDUAL DIFFERENCES ON RECALL

Degree of Doctor of Education

Institute of Education

Denise Anne Mosher

June 2018

Author's Declaration

I confirm that this is my own work and the use of all materials from other sources has been properly and fully acknowledged.

Signature:

Denise Anne Mosher

Abstract

The exponential growth of technology has placed today's educational system in a quandary, where many schools are endeavouring to meet the requirements of the current digital generation without knowing how mobile technology affects learning. Mayer's (2005) Multimedia and Redundancy Principles of Learning offer explanations for learning, and were the key theories against which the current research was measured. The aim of the investigation was to measure learning outcome in three separate subject disciplines: science–topic heart; geography–topic map; and English–topic poem, over two testing times, to determine whether mode of presentation (paper vs. M-technology, i.e., Pads®) and cognitive load (text-only vs. text & graphics vs. graphics & audio vs. text, graphics & audio) had an impact on recall. The variables of gender, working memory, and motivation were identified as possible individual differences affecting learning outcome. An opportunity sample of 346 secondary school students, males and females 11-14 years-of-age, from a multinational independent school in The Middle East participated in the quasi-experimental research. Statistical analyses included group comparisons (ANOVA, ANCOVA) with supporting correlational analysis. Prior knowledge had an impact on recall in the heart topic. Findings revealed no significant difference to learning outcome between paper and M-technology resources in each of the three curriculum topics, except in the science topic where boys' retention of information was significantly better in the paper condition than the M-technology condition. Cognitive load had an impact on recall in each of the three subject areas, where a different cognitive load combination resulted in the significant retention of information in each of the curriculum areas. Motivation and gender modulated the effects of recall. Findings across the different subject disciplines either supported or refuted Mayer's (2005) Multimedia and Redundancy Principles of Learning, determining that no one cognitive load combination was suitable for all three subject disciplines and the type of content should dictate the cognitive load condition most effective for learning. Discovery informs teaching methods and warns educationalists about making claims for innovation without any data to support gains.

Acknowledgements

I would like to dedicate my research to my father and mother, Jimmy and Lyn Erlank, who have instilled in me a love of learning; my daughters, Chantal and Bronwen, who are an eternal source of love and inspiration; and my loving husband, Brian, who has provided the emotional support and cups of coffee at 4 a.m. (even on our wedding day).

In addition, I would like to thank my wonderfully patient supervisors, Professor Daisy Powell and Professor Rhona Stainthorp, for their wisdom and guidance throughout the whole process, as well as the extremely supportive head teacher Mr Paul Young, parents, teachers, and students of the secondary school in which I conducted my research, and my study buddy Dr Terry McAdams, who shared in this journey with me.

Thank you all so very much!

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

The purpose of this study was to examine the process of learning to determine whether the mode of presentation (content presented on paper or technology media) and cognitive load (learning with different combinations of text, graphics, and audio) had an impact on recall in adolescents. This research was based on a personal interest in memory and learning, and in understanding how to support students in reaching their potential, having experienced the evolution of teaching methodology from chalkboards, to interactive white boards, to smart LED touch screens.

A means by which teachers can better understand the learning process is to ascertain how students acquire knowledge–this should not only come from reading books and journal articles, but for teachers to become actively involved in research. Wilson (2017) identified the need to connect practice to empirical evidence and for teachers, with a real connection and knowledge of what is happening in the classroom, to become researchers. As a research practitioner of psychology, sociology, and computer science, I have a genuine interest in understanding how students learn and in finding ways to improve retention and outcome, as recognised by Cromack (2008).

With the evolution and global embracing of mobile technology (M-technology) and its widespread implementation in education, research into this topic is current and relevant. A closer examination is necessary to determine exactly how successful mobile technology is for learning and whether it can be used to help improve student achievement. This concurs with Sharples, Taylor and Vavoula's (2005) and Sharples' (2013) call to formalise learning for the technology era. More research should be conducted globally, as there is limited empirical evidence from secondary schools to identify how technology could best be utilised to engage students and support learning (Pérez-Sanagustín et al., 2017). The current study has responded to this call by providing empirical evidence in an attempt to supply answers to this problem.

1.1 Knowledge, Learning, and Memory in Education

To understand the concept of learning from an educational perspective, and to understand the current research framework better (Eisner, 2017), one must first comprehend the significance of key words such as knowledge, memory, and learning that take place within an educational setting. Education has been identified as the practice whereby culture-based norms and principles are conveyed to learners who then become autonomous (Biesta, 2017).

1.1.1 Knowledge Defined

Lehrer (2018) identified knowledge as a sense of consciousness, of recognising something that is true, e.g., the sun rises in the east and sets in the west. Once this occurs a person has correct information regarding the world, and it is this information and the acceptance of it, that is "… necessary to human knowledge" (Lehrer, p. 6). Knowing that information is correct, i.e., the capacity to identify what is true, allows one then to draw conclusions, thereby leading to the ability to reason. Therefore, knowledge empowers one to realise truth and explain reality. Noddings (2018), on the other hand, recognised that knowledge is fashioned by power and is not an identified truth as identified by Lehrer. Noddings stated that knowledge is contaminated by politics and the focus should remain on the sociology of knowledge. This perspective is supported by Pritchard (2018) who identified that a fundamental change in understanding knowledge is required, one that is more than just the belief in the truth of information. Noddings identified that Dewey offered a naturalist approach, one where knowledge is based in experience–the more knowledge has been practically tested and deemed to be effective, the greater the affirmation. Dewey emphasised that knowledge required verification as opposed to justification.

Even though the explanations above differ, all have one thing in common–recognition that knowledge is greater than information alone and often includes an individual's experience.

1.1.2 Learning Defined

Learning has been linked to knowledge. As defined by the Oxford English Dictionary, learning is "knowledge acquired through study, experience, or being taught." Lowyck (2008) also recognised a connection between learning and knowledge, and identified factors necessary for the acquisition of knowledge–Lowyck proposed that "Learning is cumulative and an object of mental effort, motivation, and cognitive skilfulness… [a] gradual transformation of relevant information into knowledge." (p. xiv). Learning is, therefore, the process of knowledge acquisition which differs between people as not everyone expends the same amount of effort, motivation, and/or mental capacity. After the initial acquisition, the process becomes more complex where new information is integrated into prior knowledge, which results in increased abstraction (Graf & Kinshuk, 2008). Siemens (2005) recognised that with an increase in processing, action is required that draws on both primary and secondary knowledge, and that this ability of combining information to identify connections and patterns is an important skill. Therefore, "Learning is not so much defined by the quantity of information available but by the gradual transformation of relevant information into knowledge" (Lowyck, 2008, p. xiv).

Riding and Douglas (1993) acknowledged that learning styles are fixed and do not change within individuals. Learning styles are characterised as either the processing of information as a whole or in parts (wholist-analytic style), or representation of information in either verbal or image form (verbal-imagery style). Riding and Douglas identified that individuals tend to process information via one channel, i.e., either text or graphics, but not both. This perspective is outdated today and contrasts with Paivio (1990) who identified the dual code processing system that uses two channels, i.e., graphics (visual) and sound (auditory).

Providing a definition of learning is a precursor to understanding the term 'learning outcome.' Learning outcome is a term that has been used within the current study, but it is important to identify what the term actually means. Phillips, Porticella, Constas, and Bonney (2018) idenfied inconsistency across research in identifying predictors of learning that produced measurable learning outcomes. In many cases learning outcomes were achieved by providing an initial set of objectives against which the outcome could be measured, e.g., content knowledge, learning skills, etc. However, not all outcomes were operationalised, i.e., measurable. Phillips et al. stated that learning outcomes should be articulated, observable, and measurable as this would allow comparisons between research findings to be conducted. Brinson (2015) used the term 'learning outcome achievement' to demonstrate that learning outcome was related to attainment in tests. Conducting a meta anlysis with 59 studies in science from 2005 onwards, that compared learning outcome and methods of learning, i.e., traditional hands-on versus remote, virtual learning, Brinson found that in 89% of studies (N = 50), learning outcome achievement was better in virtual, remote studies. In addition, Brinson found that 95% of studies identified outcomes were related to content knowledge. Brinson was not alone in linking learning outcome to achievement based on content knowledge (Chen, Yang, & Hsiao, 2016; Lin, Chen, & Liu, 2017; Merchant, Goetz, Cifuentes, Keeney-Kennicutt, & Davis, 2014). Therefore, in line with previous research, learning outcome in the current study referred to achievement scores gained from content knowledge.

1.1.3 Memory Defined

A simplistic definition for memory has been provided by Werning and Cheng (2017) as "Making available information of the past for present purposes" (p. 7). Instances of memory include the recollection of events, people, and objects, both concrete and abstract. Memory is a prerequisite for knowledge, as without memory we would be dispossessed of knowledge. According to Squire and Zola-Morgan (1988), memory comprises of declarative memories (explicit episodic and semantic events which can be articulated, e.g., knowing that there are seven days in a week) and non-declarative memories (implicit and non-articulative, e.g., the skill in knowing how to ride a bicycle). The distinction between memory and knowledge has been identified by Reber, Knowlton, and Squire (1996) who acknowledged that memory is the process of creating and learning skills, and knowledge is the utilisation of those skills.

The duration of memory has been recognised by Atkinson and Shiffrin (1968) as: short-lived sensory memories that make a brief appearance before they disappear (decay); short-term memory (STM), i.e., working memory (WM), that lasts about 30 seconds; and long-term memory (LTM), where information is relatively permanent. However, this model has been criticised as artificially separating memory into short-term and long-term memories when they should be construed as the same unit with different properties (Schweppe & Rummer, 2014; Werning & Cheng, 2017). An example of a unified model of memory would be the embedded-processes model (Cowan, 1999), that unites WM, LTM, and the attention processes as a singular, integrative model of learning that involves the two systems of sensory memory and LTM.

The difference between learning and memory has been identified by Terry (2016) where, "…learning refers to acquiring knowledge or behavior, …[and] memory refers to retaining and recalling the knowledge or behavior" (p. 11). Learning has also been described as a method whereby learners interpret the world they observe through dialogue with others, which is facilitated "by knowledge and technology as instruments for productive enquiry, in a mutually supportive and dynamically changing relationship" (Sharples, Taylor, & Vavoula, 2005, p. 7), and as "…a subjective construction in which students must be fully involved" (Spiegel & Rodríguez, 2016, p. 847). Therefore, learning does not equate to just memory or the retention of factual knowledge, it is far greater–it requires active participation in the process.

1.1.4 The Importance of the Learning Process in Education

With today's emphasis on students acquiring transferable skills within schools it is important for students to gain information–not only that an event occurs, but to understand how it occurs. This knowledge can then be transferred to new situations which leads to a deeper learning (Pellegrino, 2018). For example, in the blood circulatory system the valves of the heart open and close to prevent the backward flow of blood (John Hopkins Medicine, 2018). This factual knowledge of blood flow can then be applied to different circumstances, e.g., the actuator in a car ventilation system (equivalent to the valve in the heart) controls flaps that open and close to allow conditioned air, heated to a specific temperature, to circulate (Texas Instruments, 2018). A similar process is applied within different contexts, identifying that the initial understanding of factual knowledge can support later learning. This concurs with Bruno and Dell'Aversana (2018) and Karunanayaka, Naidu, Rajendra, and Ratnayake (2017) who identified that learning occurs in context and is reflective. Knowledge is growing exponentially and the life of knowledge, i.e., how long it is relevant, can be measured in months and years (Siemens, 2005). This is a valid reason for students to acquire the skill of being able to transfer their application of knowledge from one context to another.

1.1.5 Knowledge, Learning, and Memory Summary

Within this section, discussion of knowledge, learning, and memory have been shown to be interlinked and important for the development of current learning and future application and transfer of knowledge. Yet, this would not be possible if it were not for the initial learning of basic information and being able to retain that information in memory for later use. Improving memory and the retention of factual knowledge is important to learning, as it provides the foundation on which all future learning and application and transference of knowledge takes place. The focus of the current study was on the attainment of basic knowledge which could be the foundation for a deeper learning over time. In addition, research into learning and memory has the potential of being able to devise strategies for memory enhancement for students in the future.

1.2 Identifying the Problem–Learning in Education, Pyschology, and Technology

In order to conceptualise the context of this study, it is important to determine the historic progression and evolution of the relationship between education and learning from pedagogy, cognitive psychology, and technology paradigms, to gain a holistic perspective of the development of technology, its use in education, and also to understand properly the importance placed on learning over time. The following paragraphs establish how the connection between education, psychology, and technology came about.

Prior to 1900, educational training and practice was not influenced by theory and was deemed an inexact science (Wiburg, 1995). The United States of America had become industrialised and was transformed by railway development and wealth resulting from manufacturing plants. In the latter part of the 19th century, more students needed to be educated, which was a direct result of a growing population and migration. Schools were responsible for providing a consistent curriculum, and became "…public, secular, compulsory, and free" (Handlin, 1959, p. 17). Education became political and regularised (Machin & Vignoles, 2005), and programs were implemented over an extended time frame of eight years, in the hope of creating a stable society (Handlin).

Education practice in the early 20th century motivated Dewey to explore educational pedagogy. Dewey (1929) believed that learning should be individualised and interactive. Concurrently, European scholars were questioning traditional forms of education, and theorised that it was possible to develop methods that would make the study of human behaviour more scientific (Wiburg, 1995). So, when schools faced criticism in the early 1900s of inefficient practices and poor learning by students, they turned to scientific management as a way to improve. Dewey (1929) embraced this approach and focused on applying science to educational practice. His scientific views were activity driven, and directed by the learner's sense of uncertainty when presented with new experiences and ideas. Dewey argued that the customary reinforcement of information, given to students who then memorised it, only leads to superficial learning. It is important that the student should want to learn, and actively engage in the learning process. Dewey believed the role of the teacher is to create a classroom in which the child is presented with challenging situations that he or she would be motivated to resolve.

In 1954, Skinner described the conditions of the classic classroom as unfavourable to learning, in which a single teacher was not able to support the erudition of 30 or more students appropriately. Skinner proposed improving learning via a machine that was intelligent, that could determine a level achieved, and then move on to another more complex level once an individual’s response had been analysed (Skinner, 1954, 1965). Skinner was one of the first theorists to link a technical device, i.e., technology, to learning, that was a point of interest in the current study. Concurring with Dewey, Skinner also recognised the importance of looking at individuals and acknowledging their differences: Moulded by genetics and environment, students are able to make choices and regulate their behaviour. In addition, Skinner identified interest as important for learning, but as an individual difference, as what interests one child may not interest another (Skinner, 1965). Skinner and Dewey identified individual differences, including motivation, which has also been recognised in the current study.

Running parallel to the progression occurring within education, were developments in cognitive psychology, in which a relationship between memory and learning had been established. As early as 1885, Ebbinghaus (1923/2013), while researching memory, identified that individual differences in recall were dependent on content, interest, and attention. Ebbinghaus acknowledged that repetition is important for recall, especially within extended text such as poems. Hebb (1949/2005) was one of the first scientists to identify a model of memory that linked learning to physiology, in which short-term memory (STM) was affected by electrical activity in the brain, and long-term memory (LTM) was affected by fluctuations in the nervous system. A few years later, Atkinson and Shiffrin (1968) expanded Hebb’s theory and proposed that information presented from the environment, was held in temporary STM that also served as a working memory (WM), a workspace for reasoning and comprehension, as well as long-term learning. This correlates with the current study which explored the retention of information over two testing times, i.e., a short period of time and a longer period of time.

Technology was also developing at this time and the rapid growth of audio-visual tools during the first half of the 20th century, used initially as a method for training groups in the world wars, made it possible to use sound and visualisation for the design of required learning (Gagné, 2013). The resulting training materials could then be used repeatedly, without the need for extensive teacher preparation. Gagné described early instructional technology as the confluence of the scientific study of human learning and the availability of new technologies. This perspective supported the earlier stance of Schön (1983), who noted that "as the scientific world-view gained dominance, so did the idea that human progress would be achieved by harnessing science to create technology for the achievement of human ends" (p. 31); the technologies of interest included both procedures and tools, i.e., the way things were done as well as how they were done. New techniques, such as programmed learning tied to the use of audio-visual materials, were conceptualised as a way to increase the precision with which the learner is appropriately stimulated, and thus increase learning. Such materials were easily replicable and usable in faraway locations without additional teacher training (Gagné).

With today's exponential growth of technology, traditional education systems put in place in the early 20th century are in a state of flux, unable to meet the requirements of the current digital generation (Bennett, 2012). Students have become autonomous learners within an educational setting (Fitó-Bertran, Hernández-Lara, & Serradell-López, 2013), and the aim of education has changed with regards to the delivery of content, from teachers conveying information to students while attempting to shape their character (Wiburg, 1995), to students becoming independent in their learning and teachers acting as facilitators of learning (Churchill, 2018; Gómez-Pablos, del Pozo, & Muñoz-Repiso, 2017).

This discourse has identified the development of education and learning over time, building close links between pedagogy, psychology, and technology, thereby identifying a holistic approach. As proposed by Fischer et al. (2007), "It is time for education… and cognitive science to join together to create a new science and practice of learning and development" (p. 1). This concurs with Repovš and Baddeley (2006), who recognised the advantage of a comprehensive, multidisciplinary approach to researching memory. More than one discipline needs to be consulted for learning, as without "…insight into the psychological structure and activities of the individual, the educative process will… be… arbitrary" (Dewey, 1929, p. 75). The current study adopted this approach and examined learning against the backdrop of education, psychology, and technology.

1.3 Choice of Curriculum Subjects

Research conducted in schools has often been in subjects identified by the UK government as core subjects, i.e., mathematics, English, and science. With the increased focus of technology in schools, computing is now considered a foundation subject and is a popular subject for research due to its ubiquitous nature (Gathercole et al., 2016; McMahon, Wright, Cihak, Moore, & Lamb, 2016; UK Government, 2018). This section outlines the reasons for the selection of curriculum subjects (science, geography, English) in the current study.

1.3.1 Science as a Research Subject

Coccia (2018) highlighted the importance of science by providing a global perception of science, where "Science is an organized social effort that inevitably reflects the concerns and interests of nations and society to achieve science advances and discoveries that are spread to the rest of humankind." (p.5). Supporting this statement, but within a more localised context, Saido, Siraj, Nordin, & Al Amedy (2015) identified the importance of science as a subject in education as "help[ing] students to develop their higher order thinking skills to enable them to face the challenges of daily life." (p. 13). According to Coccia, science is highly regarded in Taiwan and students achieve in this subject to gain social respect. Looking to improve learning in science, Ge, Unsworth, Wang, and Chang (2018) conducted research with 12 Taiwanese students and found the type of image used to illustrate a biological classification system impacted on reading comprehension–a coherent tree structure was more successful than the textbook illustrations. While researching a science-based subject they identified the importance of image design in pedagogy. Wang and Liou (2017) researched science and motivation in 5042 eighth-grade Taiwanese students across 150 schools and found achievement correlated positively with motivation, but differed between schools. Wang and Liou recognised the importance of researching learning within a natural environment, and that the environment itself can affect learning.

Science research in secondary schools has contributed to our knowledge of teaching and learning and provided beneficial information that has informed the methods of delivery. In addition, it has highlighted the importance of environment for learning, and has, therefore, been identified as a subject worthy of research in the current study.

1.3.2 Geography as a Research Subject

Geography is a discipline that requires spatial ability and the aptitude to read a map, i.e., a person's capacity to "perform mentally such operations as rotation, perspective change, and so forth." (National Research Council, 2006, p. 26). Maps are spatial structures that describe weather patterns, follow the spread of diseases, and can even identify geo-political strategies (Burgin, Ekström, & Dessai, 2017; Webster, 2017).

However, spatial ability and aptitude are a small part of a bigger picture, i.e., that of spatial thinking. The National Research Council (2006) identified the importance of spatial thinking: "By understanding the meanings of space, we can use its properties (e.g., dimensionality, continuity, proximity, separation) as a vehicle for structuring problems, finding answers, and expressing and communicating solutions." (p. 12). Spatial thinking can be found in different subject disciplines, e.g., parabolas in mathematics, map reading in geography, and catenary curves in physics. Spatial thinking enables patterns and relationships to be discerned and anomalies of those patterns to be detected, e.g., outliers in statistical analysis. It is important for students to develop levels of complexity in spatial thinking, for example, concrete spatial awareness of two dimentional maps can later be extended to three dimentional maps within the field of civil engineering.

Goodchild and Janelle (2010) identified that individuals who are more spatially adept have greater success in higher-level problem solving, but there is a general lack of preparation in critical spatial thinking in our education system. "The educational challenge is to teach students strategies for spatial thinking; to teach how, where, and when to use them; and to convey a critical awareness of the strengths and limitations of each strategy." (National Research Council, 2006, p. 19). The current study identified the importance of including this crutial skill as one of the topics to be researched. The current study included the spatial thinking skill of cognitively visualizing 3-D images constructed on 2-D material, one of the spatial thinking abilities identified by Lee and Bednarz (2012).

1.3.3 English as a Research Subject

Demoulin and Kolinsky (2016) linked the importance of reading and language development to working memory processes, and St Clair-Thompson and Gathercole (2006) found that working memory was closely linked with attainment in English and mathematics. Demoulin and Kolinsky's focus was to determine whether learning to read, i.e., subvocal rehearsal, had an impact on the development of verbal memory processes. This would then improve the strength of lexical representations, which would in turn improve the retention of information and recall. Demoulin and Kolinsky identified the importance of empirical research that included text-based materials to determine how learning to read forms cognition. In addition, reading allows for extended text to be researched, which has been, and still is more recently, limited to paired word recall (Bortolussi & Dixon, 2013; Epstein, Phillips, & Johnson, 1975; Krishnan, Watkins, & Bishop, 2017; Maxwell, 2018). Extending the length of text to be recalled, Goetz and Fritz (1993) researched abstract and concrete text that were four sentences (25 words) long with undergraduate students, and found that concreteness and comprehensibility were highly predictive of recall (the familiarity of text less so). They found that when a concrete sentence was read before an abstract sentence, retention was better than if an abstract sentence was read before another abstract sentence. Goetz and Fritz's findings informed a methodology of learning.

As the current research attempted to ascertain learning for students in a more realistic setting, i.e., material that students are exposed to on a daily basis, an extended text English-based curriculum topic was identified as appropriate in the current study–this would test students limits of retention and recall. Selecting a subject that was largely text-based and that utilised verbal rehearsal in the current study could identify a potential methodology for memory improvement.

1.4 Choice of Participants

The Department for Education (DfE) is the national government office in charge of the education system in England, UK. All schools are obligated to follow a national curriculum which, in secondary schools, has been divided into three key stages: Key Stage 3 (KS3) includes students 12-14 years-of-age and is a progression from Key Stage 2 in primary schools; Key Stage 4 (KS4) includes students 15-16 years-of-age; and Key Stage 5 (KS5) includes students 17-18 years-of-age (The National Curriculum, 2018). The national curriculum determines the standards all schools should follow to ensure consistency in education. However, this rule does not apply to certain types of schools, such as academies and private schools, who are not obligated to follow the national curriculum. The compulsory national curriculum in KS3 includes English, mathematics, science, history, geography, modern foreign languages, design and technology, art and design, music, physical education, citizenship and computing. Subjects for the current study were within those identified by the UK national curriculum, and selected participants represented a key stage.

Shaffer and Kipp (2013) identified the importance of distinguishing the differences in cognitive development in children during the adolescent phase, which has been identified as 11-19 years-of-age by the World Health Organisation (2015). This would include children in KS3-KS5. However, Anderson, Anderson, Northam, Jacobs, and Catroppa (2001) argued that a ‘flat developmental trajectory’ (p. 397) occurs in the executive functioning of children 11-15 years-of-age, where a stable period of development occurs. Previous research has been conducted with students in the whole of KS3 (Evans, Beauchamp, & John, 2014, music; Doharty, 2017, history; Yeoman, Bowater, & Nardi, 2016, perceptions of research in the sciences). Within the current study, students from KS3 were selected to take part as they did not have additional pressure of external GCSE examinations and comprised a larger sample group than KS4 or KS5.

1.5 Individual Differences

In line with current trends, as well as contributing to a greater understanding of the learning process, theories of learning (pedagogy, cognitive psychology, technology) were examined in the current study as a mixed paradigm against which learning was measured. Learning involves many components, e.g., genetics, motivation, and the environment, and these components, together with a student’s thoughts and experiences, shape behaviour

(Craik, 1943; Martinez, 2010). However, these influences and experiences differ between students (Oxford & Amerstorfer, 2018; Schön, 1983), which highlighted the need to address individual differences within the current study.

1.6 Overview of the Current Study

The aim of the current study was to examine learning over two testing times while manipulating the mode of presentation (paper-based medium vs. M-technology-based medium) and cognitive load (text, graphics, audio). The data collection, quantitative in nature, allowed for empirical evidence to be gained to support or refute the research questions. The ontological and epistemological paradigm of critical realism, a philosophical methodology that examines the nature, i.e., processes, of the world, was the foundation of this study (Morgan & Smircich, 1980), and the methodology used was based on activity theory (Kaptelinin & Nardi, 2006). These concepts are discussed in Chapter 2.

More specifically, the current research attempted to ascertain: whether students' learning is enhanced using M-technology (Mayer, 2008); whether cognitive load, i.e., amount of information that needs to be processed, affects learning (Mayer, 2008; Sweller, 2008); and whether individual differences play a part in recall (Gagné, 2013; Moreno & Mayer, 2010; Schön, 1983; Skinner, 1965). As this research was a situated study, the results are more likely to be generalisable than those from a laboratory. Also, it could provide insight into aspects to consider when planning a programme of study (Bryman, 2013), and allow a comparison to be made against existing and future research.

1.7 Structure of the Thesis

The thesis is organised into sections: Chapter 2 expounds the literature supporting this investigation of memory and learning. It identifies cognitive psychology and technology theories of learning, and considers theories of consolidation and forgetting. The chapter concludes with the research questions. Chapter 3 considers the methodology, identifying the sampling, testing resources, the process of data collection, ethical considerations, as well as the contribution made to the research from a small pilot study. Chapter 4 presents the findings of the mode of presentation research in terms of levels of significance, and discusses the findings against key literature. Chapter 5 presents the findings of the cognitive load research in terms of levels of significance, and discusses the findings against key literature. Chapter 6 discusses the implications of the findings and attempts to answer the research questions–suggestions are made for future research.

Chapter 2. Literature Review

The purpose of this study was to identify aspects that contribute to improved learning. The current research explored the mode effect by comparing paper and mobile technology (M-technology) resources, to determine whether learning is effected by the method of presentation. Cognitive load was also examined to determine overload, i.e., the greater the amount of information to process, the greater the demands on working memory, resulting in a higher cognitive load (Gathercole et al., 2016). The assumption would be that a lower cognitive load would be easier to process and more likely to be recalled, than a higher cognitive load. The current study explored the variables of gender, working memory, and motivation as individual differences that could have an effect on recall, relative to mode of presentation and cognitive load. Findings could contribute to the database of existing knowledge, and inform teaching methodology within similar contexts.

Learning has been described as a method whereby learners interpret the world they observe through dialogue with others, which is facilitated "by knowledge and technology as instruments for productive enquiry, in a mutually supportive and dynamically [continuously] changing relationship" (Sharples, Taylor, & Vavoula, 2005, p. 7). Spiegel and Rodríguez (2016) identified criteria that allow M-technology to become tools for learning that include personal involvement of the student, i.e., motivation, intelligence, and existing knowledge, as well as material conditions necessary for learning, i.e., access to the Internet and resources. Spiegel and Rodríguez acknowledged levels within the criteria vary between students, and stressed the importance of ascertaining achievement against individual differences, an approach that was adopted within the current study. Spiegel and Rodríguez identified cognitive psychology and technology that are key themes explored in this literature review. The importance of unifying the disciplines of education, psychology, and technology has been established (Fischer, 2006; Repovš & Baddeley, 2006), and more recently by Takooshian, Gielen, Plous, Rich, and Velayo (2016), who stressed the importance of preparing educational psychologists to work in a global community that included technology.

This chapter contains a thematic review of literature related to learning. It adopts the following structure: Theories of learning are discussed together with cognitive load; models of learning are identified; the mode of presentation section outlines theories of learning for both paper and M-technology modes of presentation, and identifies the exponential growth of technology worldwide; the individual differences section identifies how gender, working memory, and motivation may modulate, i.e., have an effect, on learning outcome; and finally, theories of consolidation and forgetting are discussed. This chapter provided the rationale against which the current study was measured.

2.1 Theories of Learning and Cognitive Load

Dewey and Skinner both identified the need to re-examine learning: Dewey (1959) from an academic and psychological experience through reflective action; and Skinner (1954) from individualised learning for every child. At the same time, i.e., mid-20th century, cognitive learning theories were evolving and soon encompassed education, neurology, psychology, and computer science (Harasim, 2012). Cognitive learning theories focus on mental processes, and correspond to technology processes, i.e., technology attempts to replicate the input, process, and output practices of the human mind (Hollnagel & Woods, 2005). According to Halpern (2013), cognitive psychology is closely related to education as it is concerned with how people "think, learn, and remember… [It is] the raw material… of thought…, … that is developed more or less fully depending on environment" (p. 18), and, as identified by Gottfredson (1997), cognitive psychology is "a … mental capability that… involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience" (p. 13). The process of manipulating information for activities such as comprehension, learning, and reasoning, has come to be known as working memory (Baddeley, 2000; Logie, 1995). Concurring, Wilhelm, Prehn-Kristensen, and Born (2012) acknowledged the plasticity of the brain and its ability to easily attain motor skills and essential knowledge about the world. Memory, therefore, is viewed as a function of the brain (Bergson, Paul, & Palmer, 2004).

Psychological explanations for learning were provided as early as 1949, when Hebb (1949/2005) proposed a two-component model of learning that consisted of short-term memory (STM) and long-term memory (LTM). Later on in 1960, Miller, Galanter, and Pribram proposed that working memory (WM) is a state, or location, in which information is temporarily held while transformation processing, i.e., learning, takes place (Richardson et al., 1996). Miller et al,'s concept was embraced by Atkinson and Shiffrin (1968) who identified a tri-component structure for memory and learning that was composed of a sensory register and short- and long-term stores. The short-term store was referred to as working memory (WM), in which selected information from the sensory register and long-term store, was processed. Building on Atkinson and Shiffrin’s (1968) WM model, Craik and Lockhart (1972) suggested a levels-of-processing approach to learning that included the characteristics of memory, capacity (amount of information that can be held in memory), encoding (the process of transferring information into a memory store), retention (the amount of information that can be recalled), and forgetting (when information can not be retrieved). Craik and Lockhart introduced the term “robust encoding”, in which a semantic orientated task has better recall than a non-semantic task– it produces a trace that is available to a larger range of retrieval cues (Lockhart, 2002; Lockhart & Craik, 1990).

Sweller (1988) recognised the effort made by WM to process data. Sweller’s (2008) cognitive load theory (CLT) identified that WM resources can be overloaded when extraneous, i.e, superfluous, activities are processed. Noyes, Garland, and Robbins (2004) defined cognitive workload as “the interaction between the demands of a task that an individual experiences and his or her ability to cope with these demands” (p. 111). According to Sweller (2008), different types of cognitive load exist: Intrinsic load is fixed and unchanging; and extraneous load is additional information that hampers learning. Germane load aids learning by producing schemas, and the amount of content within schemas helps to manage cognitive load. Abeysekera and Dawson (2015) proposed that expert learners might have more schemas than novices that allows them to manage cognitive load better and help with the integration of new material. Concurring, Veenman, Van Hout-Wolters, and Afflerbach (2006) identified that cognition, i.e., a mental process of gaining knowledge, develops over time.

The disciplines of cognitive psychology, education, and technology have close links (Harasim, 2012). Such close associations demanded that cognitive psychology be included in the current study to explain the structure, and provide an explanation of how learning takes place. Cognitive load has also been identified as important to learning and will be considered in the current study.

2.1.3 Baddeley and Hitch's working memory model

Baddeley and Hitch (1974) outlined a more complex cognitive theory that included the role of memory in problem solving, language comprehension, and learning, whereby memory represents a control system with limits of storage and processing capacities–it has even been linked to intelligence (Cowan, 2014; Halpern, 2013). This functional WM model identified three components: (a) the visuo-spatial sketch pad/loop that stores, maintains, and manipulates visual and spatial information; (b) the phonological loop that stores, maintains, and rehearses auditory speech-based information in STM and LTM, e.g., vocabulary development that is encoded and transferred to LTM through rehearsal (Repovš & Baddeley, 2006); and (c) a limited-capacity central-executive store, i.e., the controlling system in memory. According to Logie (1995), the central executive is in charge of reasoning, decision making, and managing the processes of the secondary systems, i.e., the phonological loop and visuo-spatial loop. The phonological loop holds memory traces in acoustic (sound) or phonological (word sound) form and fades within seconds (Gillon & McNeill, 2009; Repovš & Baddeley, 2006). It is a rehearsal process equivalent to unarticulated speech that is comparable to thought. The phonological loop's capacity is a reliable predictor of reading ability (Ehri et al., 2001). The role of the rehearsal process is to retrieve and re-articulate the contents held in the phonological store, thereby refreshing the memory trace. Speech enters the phonological store automatically, but Repovš and Baddeley (2006) acknowledged that information from other modalities, e.g., visual, are also able to enter the phonological store through recoding into a phonological form through articulatory rehearsal (mentally converting graphics into words and repeating it, e.g., seeing a picture of a frog and thinking of the word 'frog').

The capacity of the store is limited by the number of items that can be articulated before their memory trace fades. Baddeley and Hitch (2000) state that memory span for sentences, i.e., amount of information that can be held at one time, is approximately 16 words, compared to 6 for unrelated words. Capacity is also effected by: (a) the similarity effect, in which letters that are acoustically similar are confused, e.g., b and p; (b) the word-length effect, in which immediate memory for word sequences declines as the spoken length of words increases, and long words such as "opportunity" are more difficult to recall than "vase"; and (c) the irrelevant sound effect, in which unrelated speech is presented concurrent with, or subsequent to, research material (Logie, 1995).

The phonological loop is suited to sequential information, i.e., the recall of items in the order in which they are presented. Savill, Ellis, and Jefferies (2017) confirmed that interactions between the LTM phonological and semantic representations help stabilise the phonological trace for words, i.e., semantic binding. Items with a semantic relationship are formed more precisely than those that make no sense. Savill et al. concluded that semantic knowledge improves the strength of verbal STM, and information is, therefore, retained more accurately. According to Lecerf and de Ribaupierre (2005), the visuo-spatial sketchpad is suited to processing spatial locations (a process of short duration), as well as visual patterns (a process of longer duration).

Baddeley (2000) acknowledged that the original WM model has become outdated, as there are uncertainties such as how information from different modes, as well as information between STM and LTM, are integrated, or even where chunks are stored. Research, in which participants had to maintain auditory or visual images while performing tasks to disrupt either subsystem, identified there was nowhere else to store the images other than in the subsystems (Baddeley & Andrade, 2000). This would suggest the existence of an additional store, capable of holding complex information in a cohesive form, as a temporary representation from both subsystems as well as STM and/or LTM. Baddeley (2000) identified the episodic buffer as an additional component of WM. This component, a limited-capacity temporary backup storage system, is controlled by the central executive and can integrate information from a variety of sources across space and time (see Figure 2.1). It is capable of holding information in the form of conscious awareness, of reflecting on the information, and even manipulating and modifying it (Baddeley, 2000).

Schweppe and Rummer (2014) argued that despite the recent addition of the episodic buffer, which is the interface for STM and LTM, a disadvantage is that Baddeley’s WM model still proposed separate stores, and as valuable as Baddeley’s theory is, Repovš and Baddeley (2006) acknowledged that by allowing passive STM to become a more active system, a foundation for complex cognitive processes could be explained and open up new lines of research.

Figure 2.1. Baddeley and Hitch's (1974) Working Memory Model, Including Baddeley's (2000) Episodic Buffer

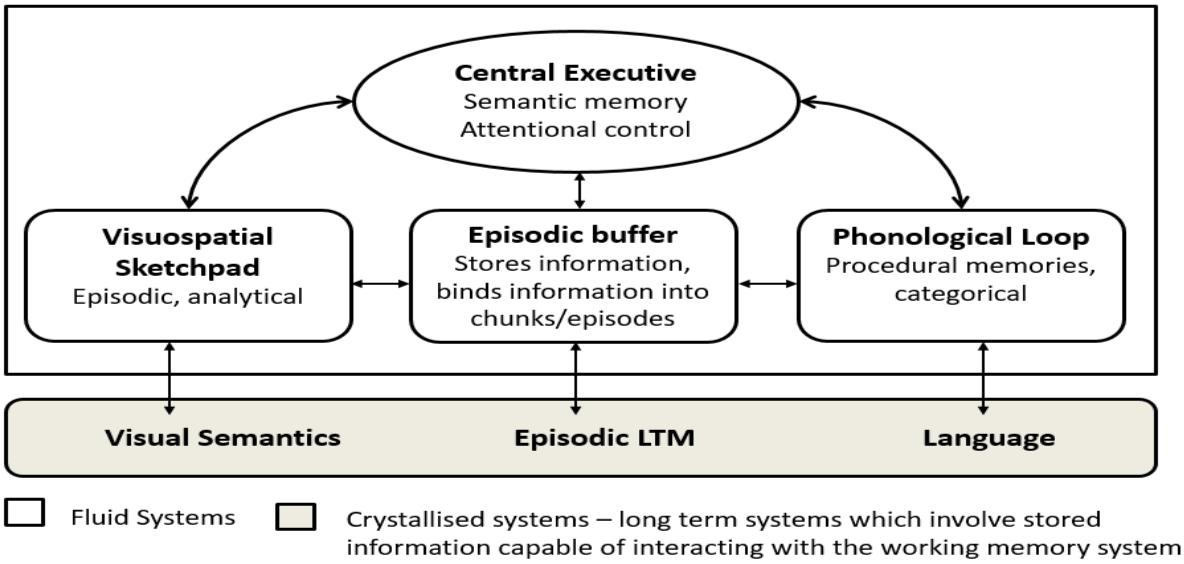


Figure 2.1. Baddeley and Hitch's (1974) WM Model, adapted from Repovš and Baddeley (2006).

Gathercole, Pickering, Ambridge, and Wearing (2004) researched the development of WM in students 4–15 years-of-age. Participants (N = 736) were randomly selected from five schools, three primary and two secondary schools, in the UK. The schools characterised average national performance in tests. The sample consisted students representing many year groups, e.g., 101 five-year-olds and 14 fourteen-year-olds. Included in the assessment were word list/nonword list recall, block recall, and the Visual Patterns Test. Findings showed a sectional structure of WM is present from 6 years onwards, and expanding functional capacity develops continuously into adolescence. However, the sample consisted of unequally represented age groups and may have provided incorrect information regarding the development of WM in children.

Baddeley and Hitch's (1974) WM model identified a structure for encoding information, and offered an explanation of how learning takes place. Gardner (1983) offered an alternative theory of learning, one of multiple intelligences that accounted for ability, talents, or mental skills that differ between individuals. Learning occurs through different intelligences and styles that enable access to understanding different types of material, i.e., logical-mathematical, linguistic, spatial, bodily-kinaesthetic, musical, interpersonal, and intrapersonal connections. Internalisation occurs with the mastery of material, and could account for learning through the use of different channels. However, White (2008) criticised Gardner by asking how one can identify whether an ability or skill is culturally important, or whether they are part of human nature. Interestingly, Gardner (2011) has acknowledged technology as a resource, but has yet to link learning styles to it.

2.1.4 Cognitive theory of multimedia learning

Even though Baddeley and Hitch's (1974) model of memory did not take technology into account, it influenced the development of a more recent theory, Mayer's (2005) cognitive theory of multimedia learning (CTML). The CTML places internal mental processes within a framework of learning with technology (Harasim, 2012; Hollnagel & Woods, 2005). Mayer proposed that learning occurs when mental representations are built from words (spoken or written) and pictures (photos/illustrations or video/animation). The learner's job, as an active participant, is to make sense of the material with the purpose of identifying new knowledge.

Mayer's (2005) model is based on three assumptions. The first is the dual-channel hypothesis, which is based on Baddeley and Hitch's (1974) WM model and Paivio's (1990) dual-processing theory, and stated that knowledge is made permanent through a verbal system dealing with language, and a nonverbal system dealing with visual objects and events. Initially, visual and auditory multimedia information are registered in a sensory memory after which attention is paid to certain items. This process of selection identifies it to WM, at which time the visual information is organised into a spatial symbol and verbal information into a verbal symbol.

The second assumption specified there is a limit to the amount of information WM can process in each subsystem–later supported by De Jong (2010) and Mayer and Estrella, (2014). Kalyuga and Liu (2015) stated that high-tech environments could result in a split-attention situation in which learners would need to distribute their attention between the material to be learnt and accessing the information. They acknowledged, however, that visual cues could reduce this effect by directing attention to salient points to be learnt. Supporting this theory, Yung and Paas (2015a) investigated the effects of an electronic character that pointed to pertinent information in a story-based animation of the circulatory system. Volunteer participants consisted of 46 fourth-grade mathematics students (M = 22, F = 24) from one school in Taiwan, and the instructional resource was aimed at teaching children logical understanding of mathematical operations in a story format. Prior knowledge of ability for mathematical calculation was determined, and scores reduced to high prior knowledge and low prior knowledge using the mean score, after which the scores were used as the covariate in statistical analysis. Cronbach's alpha for the prior knowledge test was 0.76. The testing material consisted of 20 computational problems–the testing sample group experienced pictorial representations of the question, the control group did not have the pictorial representations. Findings identified that visual representation acted as cues and helped to reduce extraneous cognitive load. Yung and Paas identified cues can be internal, e.g., colour and picture labelling, as well as external, e.g., audio voiceovers. However, it must be recognised that this research took place in a socio-cultural environment very different to the current study and the findings may not be generalisable. In addition, high and low prior knowledge was determined by means which could be misleading if there are multiple outliers. A better way to determine high and low scores would be to use the median, i.e., the most frequently occurring score.

Mayer's (2005) third assumption is active processing, whereby learners blend spatial and verbal representations together with related prior knowledge that has been activated from long-term memory. Mayer and Estrella (2014) later identified that active processing includes the selection, organising, and integration of information. Mayer's three assumptions provided the conditions for the current research that endeavoured to identify whether auditory and visual channels promote or detract from learning, and whether different types of information affect encoding in LTM.

Mayer's (2005) CTML model's structure has three stores: (a) the first allows for the perception of new information, and comprises of a visual sensory memory that holds pictures and text as visual images, and an auditory sensory memory that holds spoken words and sounds as auditory images; (b) the second WM store is where an individual processes information consciously; and (c) the third LTM store is associated with cognitive structures, and is where prior knowledge is stored. Learning occurs through a change of behaviour, such as achievement in a task.

Research within the CTML framework identified important considerations for cognitive load in education. DeLeeuw and Mayer (2008) measured the cognitive load of college students over two experiments, in which 49 were males and 106 females. Students were 17-22 years-of-age and had little knowledge of electric motors. Students watched a 6 minute narrated animation showing exactly how an electric motor works. The testing material consisted of: narration and graphics which included extraneous text with low-complexity sentences; and narration and graphics with high-complexity sentences to increase cognitive load. After watching the video, students completed a self-report scale of mental effort to determine cognitive load. DeLeeuw and Mayer found extraneous processing caused by redundant text slowed the process of learning. In addition, effort was negatively impacted by sentence complexity. They concluded that different features of cognitive load, i.e., integration, may be influenced by different types of cognitive load, i.e., level of difficulty of material presented. DeLeeuw and Mayer determined, if cognitive load was a singular-construct, the corresponding changes of all components, e.g, motivation, manipulation of testing resources, etc., should "correlate with one another." but if cognitive load consisted of multiple components that did not align, "then different manipulations of the learning situation can cause different types of cognitive load to vary." (p. 233). In this instance, "some measures are more sensitive to one type of change in cognitive load than to others." (p. 233). As a consequence, what might be measured in the analysis may not relate to the independent variables being measured and may be accounted for by extraneous processing, intrinsic processing or germane processing. DeLeeuw and Mayer's study supported the multiple-construct theory.

Moreno (2006) extended Mayer's model to take into account additional factors beyond the visual and auditory sensory inputs identified by Mayer (see Figure 2.2). Moreno's Cognitive-Affective Theory of Learning with Media (CATLM) proposed people have a limited WM capacity, and a LTM that consists of experiences and knowledge. The CATLM also recognised motivational factors affect learning by increasing or decreasing cognitive engagement, i.e., self-regulation, during which students actively monitor, adjust, and control their behaviour (Pintrich, 2003). According to Mayer and Estrella (2014), arrows from LTM are directed back to the "cognitive processes of selecting, organising, and integrating, indicating the role of motivation and metacognition in initiating, sustaining, and controlling cognitive processes during learning" (p. 14). This process is similar to the central executive's role in Baddeley and Hitch's (1974) working memory model.

Moreno and Valdez (2005) examined cognitive load and dual-channel effects. College students were presented with material explaining the causes of lightning with either text or graphics, or text and graphics. In line with Mayer's (2005) CTML, retention was highest for the text and graphics condition.

Figure 2.2. Moreno's (2006) Cognitive-Affective Theory of Learning With Media

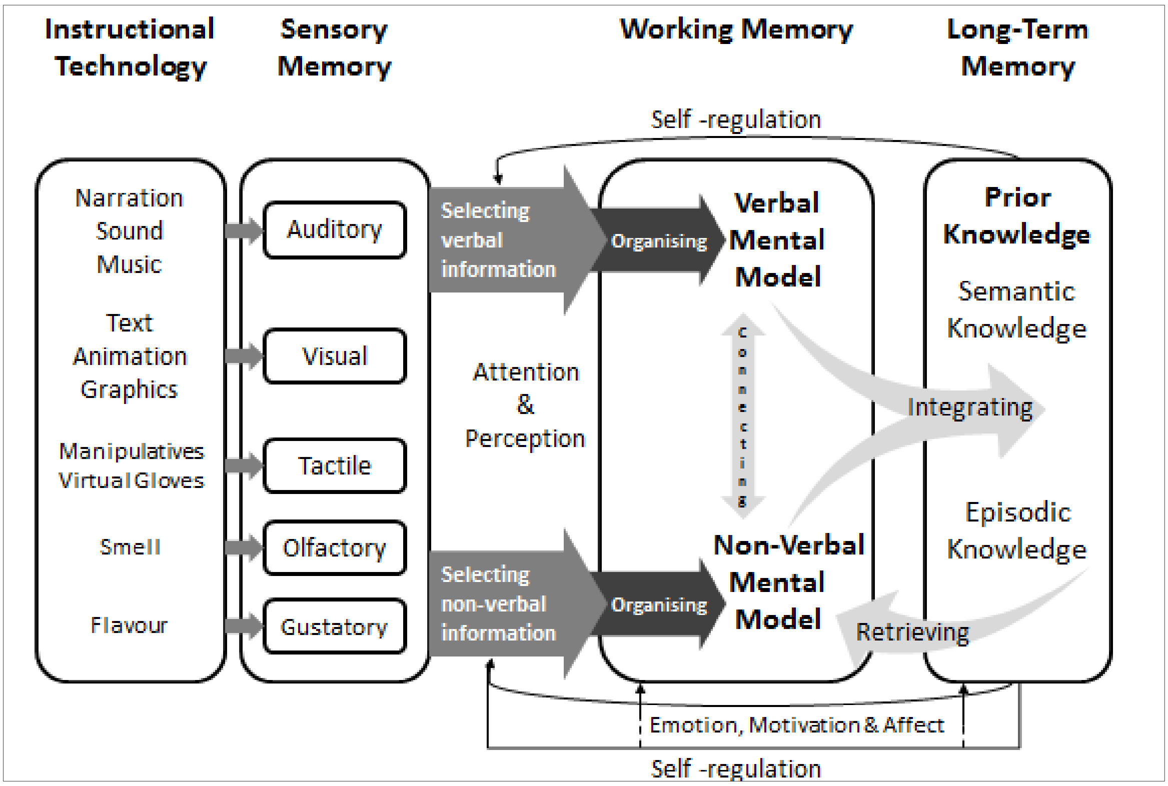


Figure 2.2. Moreno's CATLM, adapted from Moreno (2006), and Park, Plass, and Brünken (2014).

Mayer (2014) acknowledged motivation promotes learning, but only if the design does not include extraneous or distracting elements. Gillmor, Poggio, and Embretson (2015) researched whether reducing extraneous cognitive load would affect student outcome, by comparing the results of a multiple-choice geometry math assessment in which questions were adapted to remove extraneous cognitive load. Participants consisted of 222 eight-grade students from three schools and students from each school were randomly assigned to the experimental or control group. Strategies included reducing and simplifying words and graphics, and using graphics to represent spatial information. Findings confirmed reduced cognitive load improved performance. However, even though students were selected from three different schools, the sample may have been non-randomised and not reflective of the whole population. As such, the findings may not be generalisable.

Mayer and Estrella (2014) acknowledged the cognitive processing demands placed on WM capacity. Extraneous processing occurs when attention is paid to secondary items not required to be learnt, as opposed to processing the essential items. This occurs when superfluous graphics are added to resources. For example, Bunch and Lloyd (2006) identified that the use of maps in education is a key way of communicating information that is difficult to relay in words, but that graphics can hinder learning if too complex as cognitive load is exceeded.

Mayer (2005) provided a practical methodology for pedagogy and classroom practice that consisted of principles of learning. These principles included: (a) the multimedia principle in which learning is improved with combined words and pictures, than from words alone; and (b) the redundancy principle in which learning is improved with reduced input as the sensory channel can be overloaded (Mayer, 2005; Moreno, 2006). According to these two principles, multiple sensory inputs could have one of two effects, either to improve or reduce learning. The present study explored these two principles.

2.1.5 Embedded-processes model of learning

According to Schweppe and Rummer (2014), prior knowledge, which is important for learning (especially in the early stages of processing), is not fully taken into account in the CLT and the CTML. Mayer's (2005) CTML also does not account for how information is reallocated from WM to LTM. Schweppe and Rummer proposed that there are no separate systems as presented in these models of memory, and suggested an embedded-processes model of learning that included WM (the dynamic part of LTM), LTM, and attention (voluntary and involuntary).

Within the embedded-processes model, Cowan (1999) described WM as a cognitive process in which information is held in an available form. This information comprises of two systems: (a) sensory memory; and (b) information from the currently activated part of LTM, in which the level of activation determines the accessibility of information. Schweppe and Rummer (2014) suggested that the embedded-processes model of learning can be incorporated into the CTML model and provide a cohesive explanation of learning.

The embedded-processes model (Cowan, 1995, 1999) unites WM, LTM, and the attention processes as a singular, integrative model of learning that comprises of two systems (sensory memory and LTM). Cowan (1999) depicted WM as a cognitive process in which information is held in an available state. The relevant information from LTM is activated with the current focus of attention, and it is the level of activation that allows learning to take place, not the specific systems. It is possible for information to be activated, but not be in conscious awareness. It is also possible for items in LTM that are not activated, but closely associated with the focus of attention, and which could be retrieved with the correct cue in WM, to be in WM even though they reside in LTM. This infers that not all of WM is consciously experienced. There is also no additional storage system in WM for items activated from LTM, just the activation of the particular LTM information.

According to Cowan (1999), if attention is paid to a stimulus, semantic and non-semantic features are activated in LTM that results in encoding–the more features activated, the stronger the memory representation. Cowan differentiated between two methods of attention distribution: (a) altering, novel, or significant stimuli that routinely employ attention by way of an attentional orienting mechanism; and (b) the stimuli attended to by the executive control by choice, during which the representations selected are those most useful for the task. The central executive controls the focus of attention, which can be directed inwardly to LTM representations, as well as to multiple external stimuli, both visual and auditory (see Figure 2.3). This, therefore, infers that both voluntary and involuntary mechanisms determine the focus of attention.

Figure 2.3. Cowan's (1999) Embedded-Process Model of Memory

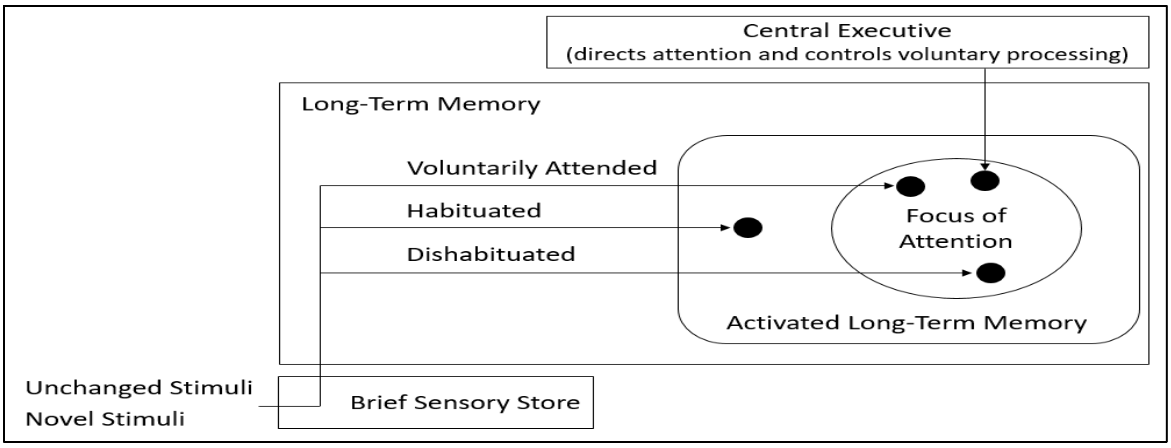


Figure 2.3. Cowan's Embedded-Process Model in Schweppe and Rummer (2014), adapted from Cowan (1999).

Empirical evidence that supports the existence of the embedded-processes model can be found in studies that identify whether STM involves LTM. Schweppe & Rummer (2014), in a discourse of learning, recognised that lists of high-frequency words have greater recall than those with low-frequency–the frequency effect suggests that words heard frequently are more accessible in the mental lexican than infrequent words.

Schweppe and Rummer (2014) acknowledged that Baddeley's (1986) central executive, which controls attention, should be incorporated into Mayer's (2005) CTML model. Also recognising the importance of the central executive, Eimer (2015) proposed that both attention and memory prioritise task-relevant material for the control of current behaviour, at the cost of less important information,. Both are limited in capacity, as attention can only be paid to a specific number of items at any given time. Furthermore, Eimer elaborated attentional systems contribute to focusing on currently perceived task-relevant representations, whereas WM controls the maintenance of the representations, even when the item is no longer present.

According to Ebbinghaus (1923/2013), activated LTM is perceived to have a time limit, which, together with decay and interference, results in forgetting. All cognitive content loses the ability to be recovered if not supported by additional learning or reviewed. However, there is a limit to the capacity of the focus of attention, which is also a contributory factor of forgetting. There is no limit to the amount of activation, but time plays an important role as reactivation processes can take place during pauses, and will counteract forgetting. Learning interference occurs when representations are very similar (Cowan, 1999). In contrast to Miller's (1956) number 7 (±2) chunks, Mathy, Chekaf, and Cowan (2018) offered that it is actually 4 (±1), unconnected items that can be held as a focus of attention at any given time, regardless of how the information was originally presented. It must be noted that as early as 2000, Cowan (2000) recognised that the STM span can hold more than four items, but this lower number takes into account the fact that the focus of attention, and activated LTM, contribute chunks to the memory span. This perspective supported Schweppe and Rummer (2014), who identified that a chunk size depends on schemas and organisation of LTM. It is, therefore, the focus of attention that has capacity limitations and not WM per se. This is in line with Baddeley (2000), who suggested that LTM is used to integrate words into chunks, and the number of chunks determine capacity.

With regards to forgetting, Cowan (1999) identified a mechanism used for maintenance. It is the role the central executive plays in reactivating items that are related to the focus of attention. This is achieved by performing a mental search through a set of objects. This process demands the use of cognitive resources that are provided by the central executive. However, encoding does not guarantee that everything encoded will be reactivated. Certain tasks will be more demanding than others, and resources will be focused on reactivating the most useful representations for a specific task. This applies to all types of representation, and not just articulatory as proposed by Baddeley. According to Cowan, the embedded-processes model does not differentiate between verbal and visuo-spatial sources of information, they are regarded simply as “two varieties of memory activation” (p. 71).

The embedded-processes model views WM and LTM as structurally interwoven, so when new data is encoded in LTM, individual components actually comprise of components previously stored in LTM. An example would be the alphabetical sounds of which words are composed, or music notes which make up songs. According to Schweppe and Rummer (2014), it is the combination of current components, new and/or previously attended to, that are stored as a new and/or different LTM trace.

Certain factors affect the ability to encode and process information differently, e.g., processing speed and experience. Therefore, individual differences could affect outcome, and needed to be considered in the current study.

2.2 Mode of Presentation and Theories of Learning

At the heart of the current research is an investigation of the way in which mode of presentation of material may impact on learning. The term ‘mode of presentation’ has been interpreted differently over time and this section outlines the different contexts in which it has been used in research.

Penney (1974) referred to the auditory and visual modes of presentation as ‘presentation modality’ and later (1989) linked modality to memory by acknowledging the existence of a modality effect in STM where the active processing of auditory and visual information occurs. Penney (1989) identified that auditorily presented tasks resulted in better retention in STM than visually presented tasks, and modality effects were not significant in LTM tasks. However, Penney acknowledged that “Not only should there be observable modality effects in long-term memory tasks, these effects will reflect the different properties of the auditory and visual perceptual and memory systems” (p. 399). Penney identified that the presentation of different types of information, i.e., auditory and/or visual, have an effect on the retention of information in STM and should also have an effect on the retention of information in LTM. This early acknowledgement of presentation modality and learning outcome was a precursor to the current study.

While researching the effect of cognitive style and mode of presentation of learning performance, Riding and Douglas (1993) referred to mode of presentation as ‘verbal descriptive information’ and ‘imagers’ (p. 306), identifying text and graphics. Mayer and Moreno (2002) identified ‘mode of presentation’ as text, graphics, and audio presented on a monitor. In 2005, Wastlund, Reinikka, Norlander, and Archer used the term ‘mode of presentation’ to represent the difference between information presented on computer monitors and paper. In 2005, Van Der Elst, Van Boxtel, Van Breukelen, and Jolles researched verbal learning with 1780 participants. Non-emotive words were presented electronically either on a computer monitor (visually) or via auditory means, a distinction that was referred to as ‘presentation mode’ or ‘the modality effect’. Findings identified that an audio version of the verbal test was better for recall from working memory (WM), while a visual version of the verbal test was better for recall from LTM, i.e., after a period of 20 minutes. Supporting the earlier findings of Penney (1989), Van Der Elst et al. identified the way information is presented has an effect on learning outcome. Schunk (2012) and Wong, Leahy, Marcus, and Sweller (2012) used the term ‘mode of presentation' to describe representations consisting of text, graphics, and audio. To avoid confusion, for the purpose of the current study, the term ‘mode of presentation’ is used as in Wastland et al. to represent learning on paper-based resources and M-technology-based resources.

2.2.1 An Overview of Technology

While Skinner (1965) was proposing an intelligent systems machine, geared to individualise learning, technology was developing through political and social channels (Jang, 2000). An example of this, is the task analysis approach of interaction between humans and machines for weapon audio-visual training for soldiers during the Second World War. Skills-based and outcome-focused, the task analysis approach was later embraced by educational institutions (Wiburg, 1995).

Subsequent to the task analysis approach, technology has developed exponentially to the point where the use of M-technology has become far more interactive, and largely universal (Motiwalla, 2007; Wingkvist & Ericsson, 2013). CISCO (2017), a multinational technology corporation based in California, confirmed that globally there are already, on average, 2.3 networked devices per capita that will increase to 3.5 by 2021, and traffic from wireless and mobile devices will account for more than 63% of IP traffic by 2021.

Smartphone traffic (33%) will surpass computer traffic (25%) before 2021. The use of tablets and smartphones is increasing exponentially with a growth rate of 29% and 49% respectively. Consumer Internet video traffic will increase from 73% in 2016 to 82% by 2021. Kabali et al. (2015) researched children in urban, low-income, minority community households, and found that at 4-years-of-age, half of the children had a television, and three quarters a mobile device. Students use mobile devices daily and are highly proficient with them (Cavus & Uzunboylu, 2009; Motiwalla, 2007; Nedungadi & Raman, 2012), and a large majority of students (96.6%) are familiar with M-technology, and even use electronic devices before their first birthday (Kabali et al., 2015). Owens (2014) and the Adolescent Sleep Working Group identified that adolescent students are often on four devices simultaneously after 9:00 p.m.

The use of M-technology in education–including content being provided in different types of media, such as images, sound, and animation–has been achieved in the 21st century. Tablets serve not only to view teaching materials, but are tools with which students can interact, draw, take notes, and watch video lectures; students now have better access to learning material (Fojtik, 2015; Hartnell-Young & Heym, 2008). Released in 2010, Apple's iPad® provides learning that can support a user's existing knowledge (Herrington, Mantei, Herrington, Olney, & Ferry, 2008; Melhuish & Falloon, 2010). Enriquez (2010) found that tablets had a substantial impact on performance, and Li (2007) discovered that there was positive student perception towards their learning experience when using technology. A meta-analysis conducted by the US Department of Education researching multimedia, found an increase in motivation, attendance, and enjoyment (Means, Toyama, Murphy, Bakia, & Jones, 2009). M-technology and mobile learning (m-learning) have the potential to become an effective teaching pedagogy in problem-solving tasks, when combined with teacher-directed learning (Cavus & Uzunboylu, 2009; El Zein et al., 2016; Motiwalla, 2007).

Global trends are impacting on the use of technology in education, and policies are evolving to provide guidance and expectations to schools (Machin & Vignoles, 2005; Yusuf, 2005). Pollard and Pollard (2004) iterated that for the past 20 years "government-funded policy reports have repeatedly identified the need for research on the effect of educational technology on teaching, learning, and schools" (p. 158). Practical examples that exemplify the implementation of these policies are the cyber charter schools in America (Ellis, 2008) and the use of tablets in many schools globally (Leinonen, Keune, Veermans & Toikkanen, 2016; Looi et al., 2009). Yet, some teachers are resistant to change, and by not embracing modern technological developments in education, influence their students to support traditional methods of learning (Tondeur, van Braak, Ertmer, & Ottenbreit-Leftwich, 2017).

On the other hand, the evolution of technology has raised certain issues, i.e., enthusiasts have tried to force the use of technology in education without proper foresight. Njenga and Fourie (2010) have coined the term “technopositivist ideology” which is defined as “a ‘compulsive enthusiasm’ about e-learning... by the people who are set to gain without giving the educators the time and opportunity to explore the dangers and rewards of e-learning on teaching and learning” (p. 199). This point of view is supported by Foulger, Amrein-Beardsley, and Toth (2011), and Emerson and MacKay (2011), who warn of the dangers of embracing technology and change when there is such little empirical evidence available. Leinonen et al. (2016) have identified an imbalance in research, with fewer studies of audio-visual in primary and secondary education than text based studies. Yet, if educational practice had to wait for empirical evidence to guide learning practice, the gap between learning and the world outside education would be greater, and as identified by Margaret Hawkins, Professor of Education, “We need to find a way to bridge the inside school-world and the outside-school world… schools are light years behind what’s out there” (as cited in Smart, 2008). It is, therefore, important to identify the value technology and paper have to learning (Pollard & Pollard, 2004).

2.2.2 M-Technology as Mode of Presentation

Looi et al. (2010) have identified that methodology has shifted over time from being instructive and teacher-led, to dynamic (continually changing) and student-centred. This progression can be seen from the theories of constructivism in the 1980s, problem-based learning in the 1990s (Compton, 2013), and high performance learning in the 2010s (Eyre, 2016). This development is not reliant on a particular medium of learning, i.e., paper or technology, but is better suited to technology that allows for greater variation of modes of learning (Herrington et al., 2008).

M-technology should be examined against theories of m-learning. Compton and Berg (2013) recognised m-learning as “learning across multiple contexts, through social and content interactions, using personal electronic devices” (In Compton, 2013, p. 4). M-learning context refers to learning independently or directed, within or outside an educational setting, during which the external environment may or may not play a part. Ally and Prieto-Blázquez (2014) acknowledged that the focus in m-learning is not on the technology itself, but is more about the learner and being able to learn in any setting. Traxler (2011) identified that m-learning provided for different types of learning, that include: (a) situated learning–that takes place within a realistic setting; (b) authentic learning–that is goal-orientated; and (c) personalised learning–that takes into account an individual's abilities and interests. The current study is interested in personalised learning. However, it could be argued that m-learning principles can be applied to both paper and technology.

Initial research examining the practice of web-based learning raised concerns regarding the use of technology. Emerson and MacKay (2011) asked whether an Internet environment, which may be associated with games and other social activities, may result in students engaging less effectively with learning resources. There is also the possibility that diverse types of interactivity may affect learning outcomes differently, and further investigation was recommended by Emerson and MacKay. This perspective supported Mehlenbacher et al. (2000), who stated that research on the complex relationship between learners and different web-based learning environments requires careful consideration in how such settings support or discourage learners, and that note should be made of learners' previous experiences and how their learning styles interact with tools intended to support student–teacher, student–student, and student–content interaction.

Billings and Mathison (2012) researched the use of iPods® with 240 Fourth-grade students, whose second language was English, from San Diego. Students completed a science task. The study was a quasi-experiment with a mixed methods design. Students were given podcasts for a five day period and asked to view one each day. Podcasts consisted of text, graphics, sound (English and Spanish) and video and lasted 4-6 minutes. A control group watch the DVD English version as a class. Billings and Mathison identified that the mode of learning is essential; students excelled when using mobile technology–this was due to learner control in choosing the frequency and duration of viewing.

Concurring, Kalyuga and Liu (2015) proposed different techniques used in a technology environment could either add to or diminish cognitive load. Therefore, design is an example of an important construct of technology environments that has been identified by researchers. The design of resources has not always taken different learner backgrounds and knowledge into account, or presented subject matter meaningfully (Höffler & Leutner, 2007). Managing the pace of material presented could reduce cognitive load caused by information transiency. Dynamic graphics, i.e., video animation, may result in learning goals being overlooked due to distraction caused by extraneous material, so irrelevant activities should be avoided (Kalyuga & Liu, 2015). Lowe (2003) proposed that presentation design should enable access to content for effective learning, and when designing for learners who are unfamiliar with the subject matter, what is salient should correspond to the subject matter. Resources often have poor design features, i.e., graphics and text, whereby learners have to integrate separate sources of information that results in a split-attention effect (Mayer & Moreno, 2003). Yet, Ayres and Paas (2007) argued that animation can promote learning if designed to highlight relevant information.

Wong et al. (2009) researched a human motor process, i.e., an origami paper-folding task in one of three narrated conditions: (a) single static–a single diagram; (b) double static–showing start and finish of the fold; and (c) animated graphics. Participants could control the resource in all conditions by moving back to a specific step or replaying a step. Participants watched the 13 stages of instruction, and immediately after, folded a square of paper to match the resource without looking at the instructional materials. Learning time was 4 mins, and testing time 4 mins 30 s, in which participants had to recreate the origami design–a Viking helmet. Results demonstrated instructional animation for motor skills is superior to static graphics. Improved performance occurred after watching an animation that included human hand movements folding the paper. Consequently, Wong et al. proposed the existence of a movement processor in WM. However, Dindar, Kabakçı Yurdakul, and İnan Dönmez (2015) investigated the use of animated versus static graphics in a computer-based English achievement test with 303 voluntary seventh-grade students from seven schools in Turkey, with a mean age of 13 years. Students had basic computer skills and were randomly assigned to groups. Testing conditions were either static graphics or animation with text. Findings determined that animation increased the response time, and did not have any significant impact on test scores. The findings supported DeLeeuw and Mayer’s (2008) multi-construct theory that described the processing of different types of cognitive load, i.e., intrinsic, extrinsic, and germane. Dindar et al. stated “keep the extrinsic load at minimum levels if the intrinsic load of the material is high. However, if the intrinsic load of the material is low, extrinsic load may not matter” (p. 158).

Leinonen et al. (2016) have identified that it is important to research audio and video as learning modalities in classrooms, as today’s generation are already using audio-visual applications, and schools are already equipped with mobile devices. The new form of digital mobile technologies, such as tablets and Smart phones, are being used as an alternative to traditional forms of computer technology. Dixon, Verenikina, Costley, and Pryor (2015) opted to use the iPad® tablet to assist learning for ASD (autism spectrum disorder) students as the tablet “appeals to the strong visual modality and preference for electronic screen media which many students with ASD exhibit” (p. 194).

M-learning lacks an evidence base of comparative studies as research and practice in m-learning are still in their infancy. Evidence from research in m-learning should be treated with caution–studies have taken place in the past five years with small numbers of participants, over short periods, using novel technology. However, a composite picture is beginning to appear that demonstrates personal ownership of technology is both motivating and disruptive. Technologies could enhance learning, but only if they enhance, and not distract from, everyday activity.

Early research has supported a preference for paper- compared to technology-based learning (Belmore, 1985; Ziefle, 1998). Yet, one must remember that early computer monitors were largely cathode ray tube with phosphorescent lighting that were hard on the eyes (Computer History Museum, 2018). It has been acknowledged, more recently, that M-technology is valuable as a mode of presentation for learning outcome–the future will lean towards M-technology as a mode of learning that can provide more varied methodology for knowledge acquisition (Keengwe, Schnellert, & Jonas, 2014). Ally and Prieto-Blázquez (2014) acknowledged that the future of m-learning will include gesture-based interaction and effective computing, an interdisciplinary approach of computer science, engineering, psychology, education, and neuroscience, in which factors that affect interaction between humans and technology are identified (Calvo, D'Mello, Gratch, & Kappas, 2014). M-learning will become a virtual learning platform based in cloud learning that will allow events to be available regardless of location. Devices will identify the physiological state of the learner, and based on the emotional state, decide what the learner should do (Ally & Prieto-Blázquez, 2014).

In summary, it must be acknowledged that the introduction of mobile devices in education has influenced teaching and learning. Buckner and Kim (2013) believed that mobile devices prepare students to become creative thinkers and problem-solvers, and learning is interactive, engaging, and can be promoted within a game-like environment. In this way, M-technology has changed the way students accrue knowledge, as well the interaction between student and teacher (Billings & Mathison, 2012). In addition, mobile-device methodology, which includes text, graphics, and audio, takes into account the student who is not suited to traditional methods of teaching (Melhuish & Falloon, 2010). Kalyuga and Liu (2015) acknowledged the importance of using different instructional methods in which students' prior knowledge "need[s] to be considered in each particular case instead of following a universal fixed set of principles or guidelines" (p. 4). The relationship between learning and technology is diverse and complex, and calls for further enquiry (Emerson & MacKay, 2011; Leinonen et al., 2016).

2.2.3 Paper as Mode of Presentation

Clariana and Wallace (2002) acknowledged that paper- and technology-based assessments (the test mode effect) would not achieve the same results, and Emerson and MacKay (2011) recognised the difficulty in determining whether paper- or technology-based learning is best, as studies have not taken into account contributing factors to outcome other than mode of delivery. With the move towards technology-based learning, Harris, Mishra, and Koehler (2009) acknowledged that education now focuses too much on technology, to the detriment of content and contextual knowledge.

Niccoli (2015) compared scores from students in a paper- or technology-based tablet condition. Students had to read an 800 word article on leadership, complete multiple choice questions for accuracy, and essay questions for comprehension. Findings showed no significant difference between paper or technology conditions in accuracy or comprehension. Yet, van der Meij and van der Meij (2014) identified a different outcome when using tutorials to explain the use of software. There were four conditions: (a) paper-only; (b) paper-based preview and video procedure; (c) video preview and paper-based procedure; and (d) video-only. There was a significant difference to learning in all conditions, but more so in the technology-based conditions, in which both the video preview and paper-based procedure, and video-only conditions outperformed the paper-only condition. Research has identified the need to examine the mode of presentation more closely to determine the impact on learning over time.

2.2.4 Research material design

Design plays an important role in the presentation of material. According to Cowan (2014), materials must be adjusted to facilitate learning in WM. The level of interaction with content is a design consideration that restricts access to extraneous components and allows a student to maintain focus on the key material (Lynch & Redpath, 2014; Saine, 2012). Formatted text, as well as horizontal and vertical spacing, can affect encoding and information retrieval as “consistent spacing helps readers to see redundancies in text and read faster, see the structure of the document as a whole, grasp its organisation” (Hartley in Jonassen, 2004b, p. 922). Hartley identified the necessity for contrast between text and its background, i.e., black text on a white background, and that text should be formatted by weight and size. In addition, text should not interfere/overlap with illustrations, and should not be justified, condensed, or stretched. Schlag and Ploetzner (2011) identified that headings and sectioning help to organise information, as well as highlight key terms. Graphics should be simple, and, taking the split-attention principle into account, text and graphics should appear as an integrated unit, which reduces needless visual searches and increases cognitive capacity. In addition, lines/arrows connect text and graphics and identify the relationship between the two. Moreno (2007) found segmenting complex visualisations reduced the level of cognitive load with undergraduate teachers learning about teaching delivery.

With regards to user interaction, Norman (2013) proposed that technology design should include visible automatic routines that are intuitive and enable an object to be manipulated, e.g., a forward button. Abeysekera and Dawson (2015) concurred and claimed that learner autonomy and pacing could improve cognitive load capacity by managing WM through interactive activities, e.g., learners manipulating material by pausing, forwarding, skipping, and rewinding. However, Vandewaetere and Clarebout, (2013) argued that learner control could use WM resources, thereby increasing cognitive load. Rey and Fischer (2013) proposed that learner control could reduce cognitive load for expert learners, but increase it for novices, creating an imbalance in learning experience.

According to Ayres and Paas (2007), many graphic animations include extraneous cognitive load, whereby resources are wasted processing irrelevant information. Attention is drawn to colour and movement that drain WM resources. Ayres and Paas argued that dynamic graphics, i.e., animation, can be less effective than static graphics because they are viewed for a much shorter time, whereas static graphics are constantly available. In dynamic graphics, learners have to remember the information after it is gone and integrate it with incoming sections of the video as well as schemas from long-term memory. Therefore, WM resources are wasted on appearance rather than learning (Yung & Paas, 2015a).

Gorin and Embretson (2006) recognised the link between cognitive psychology and psychometric theory by examining lexical models of communication and paragraph item comprehension difficulty. Encoding was more difficult for high level vocabulary and more difficult to retrieve, especially in long passages that consisted of approximately 450 words, as opposed to short passages of approximately 150 words. Clark, Nguyen, and Sweller (2011) acknowledged that complexity is a relative concept that differs from person to person, i.e., experts may not find learning particularly difficult. This supports Hartley (Hartley in Jonassen, 2004b) who proposed, with regards to computer-based measures of readability, that the longer the sentences and the more complex the vocabulary, the more difficult the text will be. However, Hartley further acknowledged that this did not take into account motivation, skills, and previously acquired knowledge.

In summary, instructional design that ignores WM may not be effective, and designing material to reduce extraneous cognitive load is fundamental for learning (Sweller, 2010). The importance of design and WM has been identified and taken into account in the current study.

2.3 Individual Differences and Theories of Learning

Cardwell (2014) defined individual differences as comprising of “people [who] differ in their genetic make-up, their life experiences, their emotional disposition, intelligence and so on” (p. 122). This definition concurs with Borghans, Golsteyn, Heckman, and Humphries (2016) who noted that students with similar IQs achieve differently, and that factors other than intelligence play a role in successful learning outcome. As such, individual differences were considered in the current research as possible factors that affect learning.

2.3.1 Gender as an individual difference

Reilly, Neumann, and Andrews (2015) identified a shift in learning outcome between genders over time. Conducting a meta-analysis using data from 1990-2011, significant gender differences in achievement were identified, with males achieving higher than females in mathematics and science. This supported research from Stoet and Geary (2013) who compared the Programme for International Student Assessment (PISA) data across nations over a 10 year period. They found boys' learning outcome was higher in mathematics than girls, and girls were better at reading. A reason for this has been provided by McGinnis and Tippins (2001), who identified that the methodology of science lessons in the past has been to support boys more than girls. From within the core subjects of mathematics, English, and science, strategies such as making science relevant and useful to all students has been implemented to encourage more girls to pursue the sciences (Krapohl et al., 2014; Labrie et al., 2003). With possible gender differences existing, the individual difference of gender was considered within the current study as influencing on learning outcome.

2.3.1.1 Gender and mode of presentation

With technology becoming ubiquitous, very few studies have compared mode of presentation (paper versus M-technology) and gender with regards to learning outcome in secondary education. Studies have concentrated largely on comparing how the different genders perform on M-technology resources. Papastergiou (2009) found, despite boys' preference and experience in gaming, that learning for both boys and girls in secondary school computer science classes did not differ significantly when using technology. Concurring, Wang, Wu, and Wang (2009), found no gender differences in m-learning, and identified a strategy for promoting m-learning by encouraging users to recognise m-learning resources as playful and valuable.

A review of the published literature did not identify many studies of gender differences and a mode effect at secondary school level, but trends in practice have identified technology as a methodology for the future (CISCO, 2017; Wingkvist & Ericsson, 2013). The current study endeavoured to identify whether any gender differences existed in learning outcome between boys and girls on paper-based and M-technology-based mode of presentation, while taking into account the complexity of processing within each mode.

2.3.1.2 Gender and cognitive load

Bevilacqua (2016) identified the contribution that evolutionary educational psychology is making to explanations for gender differences. The argument is that WM processes socially germane information for each gender, i.e., visuo-spatial for males and language for females. This has occurred through evolutionary social pressure and makes primary knowledge gender-specific.

Robert and Savoie (2006) investigated accuracy and gender differences using WM tasks involving either verbal or visuo-spatial information (single span), and verbal and visuo-spatial information (double-span). Participants consisted of 50 male and 50 female undergraduate students between the age of 19 -25 from the University of Montreal, who were paid US$14.00 for taking part in the study. Testing conditions included individual visuospatial WM tasks (including Corsi’s block-tapping test) and a verbal WM task where students had to recite a series of digits in the same order in which they heard them. Students did not have a time limit, but were asked to work as quickly as they could. A male and female experimenter each examined the same number of students, and counterbalancing was implemented, where some students started with test 1 and some with test 2. Testing sessions took place on the same time each day. There were no significant gender differences in the single span task, but females achieved significantly better than males in the double-span task, i.e., females achieved better than males in a higher cognitive load condition.

Hyde (2007) conducted a meta-analysis of 46 studies to examine the effect size of gender differences, and proposed the gender similarities hypothesis that stated males and females are mostly similar. Harness, Jacot, Scherf, White, and Warnick (2008) researched modal-specific elements of WM, using text and graphics as stimuli, and found no significant difference between males and females in a verbal WM task, but when a distraction component was introduced, females performed significantly lower than males. Yet, females showed significantly higher recall than males on a visual WM task. Findings support a gender difference in cognitive ability that is contradictory to Hyde’s gender similarity hypothesis.

In summary, empirical evidence has identified the importance of considering gender differences within research. Halpern (2013) warned that there are no simple answers when researching the complex topic of cognitive gender differences, and that research results have often been misused. Variability and overlap occur within groups, and group averages need to be used very carefully when making decisions about individuals. No-one is typical of a perceived group norm, i.e., boys' spatial ability or girls' reading ability. It is important to consider gender differences to improve understanding and identify legitimate variances. This can only be done through careful research, and as such, the issue of gender differences and cognitive load was examined in the current study.

2.3.2 Working memory as an individual difference

Herrnstein and Murray (1994) acknowledged that intelligence is "something real… that varies from person to person" (p. 1), thereby exemplifying the difficulty in providing a finite definition of intelligence. Sternberg (1985) provided a social-psychometric definition of intelligence by identifying differentiated processes involved in cognitive ability, whereby "[intelligence is] a mental activity directed toward purposive adaptation to, and selection and shaping of, real-world environments relevant to one's life" (p. 45). Fischer et al. (2007) criticised Sternberg's definition by stating psychometric tests do not identify how capable test-takers are at shaping their environment as many people interpret questions differently. It also does not explain why people may choose not to spend time on questions, which could ultimately lead to an intelligent person being perceived as non-intelligent. Yet, Spearman (1927) noted the reliability and consistency of results for mental tests, where, if a person took two different tests, they would achieve similarly high or low in both. Spearman believed that learning outcome is influenced by intelligence, and that intelligence is a measure of a person's capacity for complex mental work (Herrnstein & Murray, 1994).

Herrnstein and Murray (1994) proposed that intelligence can be improved. Svendsen (1983) found former slow learners, with multiple life problems during childhood, who were tested as a child and again when 30 years old, had larger IQ gains (18.1 points) than that of learners with one or no problem who experienced no further education (5.0 points). Empirical evidence, therefore, supports the argument that the environment can affect cognition, and that it is not fixed, but fluid.

With the use of multimedia in education, De Bruyckere, Kirschner, and Hulshof (2016) discussed the general fear that technology is the cause of reduced intelligence in society and education, and proposed that the use of technology is an important contributory factor for increasing the average IQ. This perspective supports Johnson (2006), who identified the sleeper curve effect, in which the most perceived forms of mass diversion, e.g., video games and violent television shows, actually encourage a more intellectual culture. Johnson believed that technology enhances our cognitive faculties, not reduces them. These views were supported by Greenfield (2014), who recognised the complex cognitive skills required to play electronic games, in which the interactivity of technology encourages students to create information, not just utilise it.

Research has linked intelligence to WM and motivation: (a) Hoerig, David, and D'Amato (2002) found intelligence and memory to be significantly correlated with students with learning disabilities; (b) Hill, Foster, Sofko, Elliott, and Shelton (2016) found motivated individuals, with average or above average WM ability, performed better on intelligence tests; and (c) Tourva, Spanoudis, and Demetriou (2016) found WM to be the main cognitive process underlying all forms of intelligence in students, i.e., general; fluid–abstract reasoning in novel situations; and crystallized–existing skills, knowledge and comprehension. Gathercole, Lamont, and Alloway (2006) suggested that a close relationship exists between students' performance at school and their WM, e.g., low WM scores equate to poor achievement. In the UK, Gathercole et al. (2016) investigated WM, reading, and mathematics in 230 students 5-15 years-of-age. The Automated Working Memory Assessment (AWMA) and Wechsler Individual Achievement Test II (WIAT-II), used to assess learning, demonstrated that WM capacity is correlated with learning outcome, and that poor WM capacity does, in fact, impair learning. Performance in WM is associated with comprehension, literacy, mathematics, science, and English, and affects academic performance in students 7-14 years-of-age (St Clair-Thompson & Gathercole, 2006). Students who achieve low WM scores perform lower than the expected standard in English, mathematics, and science national curriculum assessments. WM, therefore, was an important consideration in the current study, and was used to identify individual differences in cognitive ability.

2.3.2.1 Working memory and mode of presentation

Andrews and Haythornthwaite (2007) identified the "changing landscape of the learning field" (p. 5), and the concern of amalgamating technology with traditional methodology. Developing technology methodologies include video-based materials, gamification for learning, audio recordings, and adaptive learning systems (Ally & Prieto-Blázquez, 2014; Andrews & Haythornthwaite, 2007). Learning material has changed from being heavily text- and paper-based, to a multi-modal technology-based approach. Andrews and Haythornthwaite recognised potential issues regarding the number of interacting elements, such as technologies and practice, when compared to paper-based learning. The move from paper- to technology-based learning seems to be natural given the global technological climate (Motiwalla, 2007; Wingkvist & Ericsson, 2013). Yet, questions need to be raised as to how this move relates to learning and WM. Is WM enhanced through the use of technology, or is learning better with the traditional paper-based medium?

Jabr (2013) accounted for the preference of paper-based text as it supports intuitive navigation, i.e., mapping the location of text on the page while reading. Monitors interfere with intuitive navigation, which could impact WM capacity. This is supported by empirical evidence from Mangen, Walgermo, and Brønnick (2013), who researched secondary students reading extended text (1,400-2,000 words) on paper and on a computer. Findings revealed that students in the paper condition demonstrated better comprehension than students in the computer condition.

In summary, WM models, as identified by Baddeley and Hitch (1974), Mayer (2005), and Moreno (2006), place importance on WM as a key component in supporting the acquisition and advancement of learning and educational skills (Lucid Recall, 2015; St Clair-Thompson & Holmes, 2008). Existing research has identified the need to compare outcome between paper- and technology-based material, but very little research exists that links it directly to WM. The current study recognised the importance of researching the mode of presentation in relation to WM, to determine the best methodology for students with high or low WM, to assess the necessity for intervention in specific subjects.

2.3.2.2 Working memory and cognitive load

The relationship between WM and cognitive load can be explained through Sweller et al.'s (1998) cognitive load theory (CLT), Mayer's cognitive theory of multimedia learning (CTML), and Paivio's (1990, 2007) dual-coding theory. The CLT is concerned with processing challenging tasks, and the focus is to control the cognitive load levied by these complex tasks (Sweller, 1988; Sweller et al., 1998). Paas, Van Gog, and Sweller (2010) and Cowan (2000) stated that unrehearsed information, held for 30 seconds in WM, together with a capacity of 4 (±1) items, places limitations on WM. Yet, these restrictions only apply to new information. When previously acquired knowledge is integrated with new knowledge, the limitations disappear. This perspective is supported by Ayres and Pass (2007), who stated that schemas are knowledge stored in LTM and afford multiple components to be viewed as a single component during processing, thereby increasing WM capacity and allowing for more intricate learning. Brady, Störmer, and Alvarez (2016) proposed that WM fills up in 100 ms, however, this is not when including existing knowledge.

Sweller et al. (1998) identified three types of cognitive load: (a) intrinsic–cognitive load that is influenced by the difficulty of resources to be learnt; (b) germane–the WM resources necessary to process intrinsic load; and (c) extraneous–the cognitive load caused by poorly designed material, during which non-essential information must be processed. The purpose of the CLT is to find ways to reduce extraneous cognitive load. Researching mathematical problem-solving tasks, Sweller (1988) recognised that cognitive effort alone did not enhance learning outcome, but paying attention to certain elements allowed learners to manage and manipulate cognitive resources.

Mayer's CTML theory (2005) also recognised the importance of WM. The model is based on the assumptions of a dual-coding system of visual and auditory channels, to which there is a limit to the amount of information WM can process in each subsystem. During the process of learning, the learner integrates the visual and auditory representations together with related prior knowledge from long-term memory (Mayer, 2009), which could overload WM and result in the obstruction of learning.

WM capacity is influenced by the complexity of information to be encoded, and simplified information can reduce the effects of cognitive load (Sweller, 1994, 2008). Sweller (2008) also identified several factors that affect the WM's ability to encode to long-term memory: (a) the split-attention effect, in which information from multiple sources decrease WM capacity; (b) the modality effect, in which integrated visual and auditory modes presented together is better than just the visual mode; (c) the redundancy effect, in which duplicated data presented via different modes reduces learning; and (d) the expertise reversal effect, when initial cueing that serves to promote novice learning reduces effectiveness with expert learners. Paas et al. (2010) extended the CLT to include vicarious learning, which occurs when students learn from the experience of others. Therefore, resource presentation, knowledge of how material should be structured, as well as activities performed, play a key role in CLT research.

Paivio's (1990) dual-coding theory included two cognitive subsystems, i.e., verbal for language, and nonverbal graphics for non-linguistic objects. As the representations are linked to sensory input, they can be viewed as multi-modal (visual, auditory) that can work separately, or combined, to facilitate verbal and nonverbal behaviour. If students have to divide their attention between different types of input, such as text and graphics, they have to process two sources of information that increases cognitive load. This supports the CLT's split attention effect. On the other hand, the use of two channels for input, e.g., audio and graphics, may increase WM capacity and reduce cognitive load caused by split attention as WM capacity can be enhanced by using two separate channels. This is known as the instructional modality effect (Kalyuga, Chandler, & Sweller, 1999).

The assumption that the combination of illustrations and text support deeper learning was tested by Yung and Paas with primary school students in a mathematics task (2015b). Results indicated that combined text and graphics resulted in a higher learning outcome and lower cognitive load than learning with text-only. A visual representation permitted the focus to be directed to the essential components, resulting in reduced extraneous load that allowed WM to build a complete mental picture of the components to be processed.

Empirical evidence that supports Sweller et al.'s (1998) CLT and Paivio's (1990) dual-coding theory, was conducted by Kalyuga et al. (1999) found WM capacity improved with the combination of audio and graphics (two channels) as opposed to text and graphics (one visual channel). Kalyuga et al. discovered multiple modalities could reduce cognitive load. However, the combination of text, graphics and audio resulted in cognitive overload that impacted learning negatively.

In summary, it has been established that overloading WM can prevent learning (Kalyuga et al., 1999; Mayer, 2005; Paivio, 1990), and theory and empirical evidence have acknowledged the need to consider the limitations of WM when creating learning resources (Sweller et al., 1998). It has also been recognised that existing knowledge needs to be taken into account (Mayer, 2009; Paas et al., 2010). In addition, research conducted by Kalyuga et al. (1999) has identified the benefits to learning when using two channels (audio-visual) as opposed to one channel (audio or visual), which directed the current study to examine learning via different channels to assess the best cognitive load combination for WM capacity.

2.3.3 Motivation as an individual difference

According to Mayer and Estrella (2014), motivation refers to “the learner’s cognitive state that initiates, energizes, and maintains goal directed behaviour” (p. 12). Kress, Sharon, and Bassan (1981) found motivation consists of choice and effort that culminates in the decision whether or not to continue any actions. Nakamura and Csikszentmihalyi’s (2014) flow theory describes motivation as an optimal-intrinsic-inherent experience, characterised by a fulfilled state of awareness accompanied by strong concentration, effortless control, and profound enjoyment. The activity is reward in itself and does not consider any extrinsic good that might result. However, Martens, Gulikers, and Bastiaens (2004), while investigating 33 undergraduate students in a simulation after which participants had to assess and write a report regarding absences for a bus company, found no differences in outcome and that students with high intrinsic (inherent) motivation do not work harder, they just have a tendency to do things differently, i.e., they have greater curiosity and explore more.

Cole, Feild, and Harris (2004) proposed that even though intelligence and ability affect outcome, it is motivation that decides the amount of effort given to a learning activity. According to the self-determination theory (Deci & Ryan, 1985), motivation increases when students, acting as autonomous agents, are actively engaged in the learning process. In addition, the type of motivation is central to engagement–intrinsic motivation occurs when activities are interesting, and integrated regulation exists when students have identified and integrated the values related to what needs to be learnt. Extrinsic motivation occurs when a discrete outcome, i.e., an external locus, such as a reward, is offered (Abeysekera & Dawson, 2015). Deci and Ryan (2000a) acknowledged that autonomy and competence are better motivators than compelling students to complete work, as being forced to complete work could lead to feelings of resentment.

Nakamura and Csikszentmihalyi (2014) regarded motivation as an intrinsic value, which concurred with the earlier research of Hickey (1997), who found extrinsic motivation has a negative impact on learning. Students with great intrinsic motivation often outperform students with little intrinsic motivation (Martens et al., 2004). Yet, this is in contrast to Cassidy and Lynn (1991) who found that materialism was a significant indicator for motivation. Hart, Stasson, Mahoney, and Story (2007) examined the relationship between five personality traits and achievement motivation. Correlation analysis indicated that conscientiousness, openness, and extraversion were positively associated with intrinsic achievement motivation, whereas extraversion, conscientiousness, and neuroticism were positively related to extrinsic achievement motivation. Results suggested that both forms of motivation may be more complex than originally expected. Expanding on this concept in their cognitive evaluation theory, Ryan and Deci (2000b) identified the factors of sense of relatedness, control, and competence as predictors of intrinsic motivation. They proposed that extrinsic and intrinsic motivation are not distinct categories, and that there are stages in motivation that vary from amotivation to intrinsic motivation.

Further reasons for motivation were identified by Hickey (1997), who examined existing research on cognitive activity and found motivation was largely influenced by interest and goal orientation. Students' belief in their own ability facilitate achievement-oriented conduct and awareness. Cassidy and Lynn (1991) conducted a longitudinal study over 7 years with 199 males and 252 females to ascertain predictors of motivation for achievement. The questionnaire was composed of 20 items. Results indicated that school-type, i.e., small classes which equate to less crowding, IQ, and home background are considered important factors for learning and achievement. Competitiveness and status aspiration were additional significant indicators that motivated individuals.

Mayer and Estrella (2014) linked motivation and the emotional design of resources (selection, organisation, integration), to the theories of CTML and CATLM. Visually appealing resources attract attention and help to sustain cognitive processes by encouraging (motivating) the learner to persevere with the process of understanding. A learning process is initiated that results in improved outcome. Mayer and Estrella conducted research using a biological comprehension test that outlined how viruses cause colds. Black and white graphics with printed text were compared to colourful graphics portrayed with human-like characteristics and emotions. Findings identified that emotional design, using colour and humanisation, caused learners to exert more effort (motivation) to understand the material, which led to improved learning. There was a significant difference to learning outcome.

Self-regulation has been identified as a key component for learning. Credé and Phillips (2011) identified meta-cognition, motivation, and behaviour as components of self-regulation, and, in turn, indicators of academic performance. Credé and Phillips proposed self-regulated learning can explain individual differences as it occurs between people and within people, which could explain why students achieve differently in the same subject, or why one student achieves differently in different subjects. This social-cognitive perspective supports García and Pintrich (1995) who stated that "research… has… shown that positive motivational beliefs such as perceptions of high self-efficacy, …mastery goals, …interest in the task or content, and low levels of test anxiety are positively related to greater cognitive engagement... [and] …academic performance" (p. 2). Comparing motivation in different subjects, Wolters and Pintrich (1998) found no difference in motivation between mathematics, English, and social studies. This indicated subjects, per se, are not necessarily a contributing factor for motivation.

2.3.3.1 Motivation and mode of presentation

In a biology River City study, Dede, Clarke, Ketelhut, Nelson, and Bowman (2005, April) found students more than doubled their outcome when using technology-based resources compared to paper-based resources. They found that students were highly engaged in the technology condition. Similarly, comparing the effect of mode of presentation, Driessen, Muijtjens, Van Tartwijk, and Van Der Vleuten (2007) examined undergraduate medical students in paper- versus Internet-based learning and motivation conditions. Results showed no significant mode effect, but a significant motivation effect, in which Internet-based work enhanced student motivation. Students achieved higher scores, and spent more time on their work, when using technology. This research is supported by Papastergiou (2009), who found no difference in outcome between males and females for digital game-based learning in a computer science class in secondary school. The game was found to be equally motivational for both males and females.

Papastergiou concluded that computer games provide a motivational learning environment.

In summary, research has identified trends in which technology has enhanced engagement and motivation (Dede et al., 2005, April; Driessen et al., 2007; Papastergiou, 2009). Prior research has identified the need to examine the factors of mode of presentation and motivation in relation to performance, which the current study endeavoured to do.

2.3.3.2 Motivation and cognitive load

Findings from existing research has identified the role of motivation in relation to cognitive load and learning outcome. However, Iten and Petko (2016) researched electronic games promoting media competency with primary school students, and found a significant relationship between enjoyment and motivation, but no significant outcome to learning. An explanation given was fun elements introduce extraneous cognitive load and prevent students from engaging with the content. Lin, Atkinson, Savenye, and Nelson (2016) researched cognitive load and intrinsic motivation against outcome with undergraduate students in a computerised interactive human cardiovascular task. Findings indicated that cueing, i.e., arrows pointing to relevant areas, reduced extraneous load which resulted in intrinsic motivation and improved learning.

Baylor, Ryu, and Shen (2003, June) explored the effects of audio and animation on motivation and learning with undergraduate students. An agent's voice, which was either human or computer-generated, was presented with or without animation. Findings indicated learning was greater when animation was present, a higher cognitive load that induces a split-attention effect (Homer, Plass, & Blake, 2008), even though participants were less motivated in this condition. Participants were more motivated in the no animation and human voice condition, which they found more engaging. Baylor, Ryu, and Shen's findings do not support Moreno's (2010) cognitive-affective theory of multi-media learning, where, if students believe they will do well and feel positive about the learning task, cognitive resources will be assigned to the task and learning will take place. Their findings indicated learning can take place even if motivation is not present. Moreno stated that learning takes place with motivation, but did not dispute the fact that learning can take place without it. Research has shown that cognitive load should be examined alongside motivation as a factor for improved outcome.

Sweller et al. (1998) identified the CLT as demands made on the cognitive system that has a limited capacity, and these demands consist of mental load, i.e., tasks and effort made by the learner. Van Merriënboer and Sweller (2005) recognised that the CLT focuses on promoting methodology that decreases extraneous cognitive load to allow for greater learning. However, Moreno (2010) identified individual difference issues with the CLT when establishing a measurement of extraneous and germane cognitive load. Distinguishing between these two types of load, and demonstrating different learning outcomes for each, is difficult due to the ambiguous nature of these varieties of load–each student is unique and cognitive load will, therefore, differ for each student. Moreno stated that the CLT does not identify a relationship between load and motivation, and that it is the level of difficulty that affects the motivation to learn, i.e., whether the task is too difficult or too easy.

Chang, Liang, Chou, and Lin (2017) researched university students taking an education course, and examined cognitive load within game- and non-game-based learning. Both groups experienced online resources: the control group, web-page resources; and the experimental group, an electronic game. The game-based participants, who experienced lower extraneous cognitive load, were more motivated and able to control their learning than their counterparts who experience higher germane cognitive load. However, Hawlitschek and Joeckel (2017) researched motivation with secondary school students playing an electronic game, and findings failed to identify a relationship between effort and outcome. Hawlitschek and Joeckel failed to find an effect on intrinsic motivation, but did find a positive effect of extraneous cognitive load.

Motivation and cognitive ability can vary between tasks, and the need for a reliable measure of self-regulation was identified by Pintrich, Smith, Garcia, and McKeachie (1991) who developed the Motivated Strategies for Learning Questionnaire (MSLQ). Pintrich (2003) proposed motivational factors affect learning by increasing or decreasing cognitive engagement, and the purpose of the questionnaire was to assess motivation and learning strategies using valid and reliable measures. Pintrich et al.'s final questionnaire was composed of 81 questions on a 7-point Likert scale, varying from "not at all true of me" to "very true of me." Tabatabaei, Ahadi, Bani-Jamali, Bahrami, and Khamesan (2017) used the MSLQ to determine low motivation students, and after teaching motivated strategies found a positive correlation between metacognition and performance.

In summary, arguments supporting or refuting the relationship between motivation and cognitive load for improved learning are complex. Studies show that extraneous cognitive load is a deterrent for learning and motivation (Chang et al., 2017; Iten & Petko, 2016), and that learning can take place without motivation (Lin et al., 2016). Further research needs to be conducted in this area to clarify any interaction between motivation, cognitive load, and performance.

2.4 Theories of Consolidation and Forgetting

Stickgold and Walker (2007) discussed the complexities of the term “memory”, stating that it covered a wide variety of memory types as well as storing different types of information, e.g., declarative, procedural, and emotional memory. Stickgold and Walker defined the term memory consolidation as “a poorly defined set of processes which take an initial unstable memory representation and convert it into a form that is both more stable and more effective” (p. 331). This process is automatic and occurs without awareness. Theories of encoding in LTM through rehearsal and reactivation have been identified (Baddeley & Hitch, 1974; Cowan, 1999; Mayer, 2005), but understanding how we retain information over a longer period of time, i.e., consolidation, should be considered.

2.4.1 Reflection as consolidation

Reflection has been identified as a cognitive process in which “reasoning, thinking, reviewing, problem-solving, inquiry, reflective judgement, reflective thinking, critical reflection, reflective practice [take place]” (Moon, 2013, p. viii). Reflection has been identified as vital to education, due to the disparity between student understanding of natural events and scientific explanations (Kori, Pedaste, Leujen & Mäeots, 2014). Students learn new concepts and make connections between knowledge and their own reality, a process that is enabled through reflection (Procee, 2006). Reflection is viewed as a process that leads to more profound learning (Kori et al., 2014) and could account for the consolidation of learning in the present study.

Boud and Walker (1998) believed reflection should have personal emotional involvement. They criticised the recognition that has been awarded to reflection and identify the poor practice of reflection in education. Yet, the skill of reflection can be developed through questioning techniques that guide reflection and creative thinking about problem-solving (Leijen, Valtna, Leijen, & Pedaste, 2012; Moon, 2013). Fiorella and Mayer (2012) found that directed prompts provided on paper gave a deeper understanding of subject matter. However, Furberg (2009) noted that predefined guidance may not be suited to every individual.

Technology environments in education are viewed as instructional systems through which students acquire skills (Wang & Hannafin, 2005) and can, therefore, be used to enable reflection (Leijen et al., 2012; Procee, 2006). Supporting this point of view, Kori et al. (2014) conducted a meta-analysis of 21 articles published between 2007-2012, and identified that within an educational setting, at times with guiding questions given in advance, the use of technology supports reflection.

Leinonen et al. (2016) researched mobile apps that utilised audio-visual modalities for primary and secondary school students. The interest was in student-centred learning, in which the use of technology was measured against the framework of Fleck and Fitzpatrick's (2010) five levels of critical reflection in human-computer interaction. Three of the five levels of reflection were: (a) level 1–no reflection occurs; (b) level 3–dialogical reflection ensures new points of view; and (c) level 4–transformative reflection that results in a change of practice. Leinonen et al. aimed to determine whether individualised learning encompassed abstract thought and the bigger picture thinking. Qualitative data were collected via interviews and participants included teachers and students in 13 European countries. Results indicated, together with guided use from the teachers, that the mobile apps allowed for more reflection. Without teacher guidance, only the first three of Fleck and Fitzpatrick's levels of reflection could be achieved. Leinonen et al.'s research confirmed that reflection can improve when combined with human interaction.

2.4.2 Theory of spaced learning as consolidation

Spaced learning is a process that requires students to rehearse information on retrieval at different points of time, e.g., homework and tests. Subrahmanyam (2017) identified that spaced learning is based on the principle of the psychological spacing effect, which is "studying via mass presentations of information yields substantially inferior results when compared to spaced presentations of the same information" (p. 4). Subrahmanyam researched long-term retention within education and identified the importance of revisiting newly taught concepts over time. The process was a programmed algorithm that would take the input (newly learnt concepts) from a student. These newly learnt concepts were inputted into the algorithm, together with a numerical representation of understanding, and the algorithm calculated a set of concepts to review daily. Students had to re-evaluate how well they understood the concept before and after revisiting the information. This new data were then entered into the algorithm to recalculate when next to revisit the concept. Subrahmanyam found spaced learning to be a highly effective way to retain information.

Boettcher et al., (2018) acknowledged that spaced learning is effective for both traditional learning and motor skill acquisition. Research with medical students learning how to tie surgeon's knots, found that the students who had trained using the space learning method produced far better quality knots than the students who learnt the more traditional way. In addition, students were not as anxious and felt more in control. Spaced learning appears to be an effective method of consolidation and retention.

2.4.3 Theories of forgetting

In addition to a lack of rehearsal and weak encoding identified in cognitive theories of learning (Baddely & Hitch, 1974; Mayer, 2005; Cowan, 1999), Mandler (1980) identified two processes necessary for the recognition of existing knowledge–both processes work concurrently. The first process identifies whether the item is familiar, and the second attempts to search and retrieve the item within the context it was originally presented. Lack of retrievability is caused by the absence of relevant cues, and the lack of familiarity by not recognising features of an item (Mandler). Retrievability, therefore, relies on the strategies involved when initially encoding data, which can differ between individuals.

This perspective was extended by Bower (1981), who acknowledged that cue availability is dependent on the consistent emotional state of a person, i.e., the mood-congruity effect, in which participants engage more with events that are consistent with their emotional state. Bower argued retrieval is more successful when a participant's mood is consistent at the time of learning the information and time of recall. Spear and Riccio (1994) proposed that changes in context affects the accessibility of retrieval cues, i.e., when cues present during the initial presentation of content are not present during recall. When the emotional state of a person differs between input and recall, the more difficult it is to retrieve information.

According to Ecke (2004), language attrition–a slowing down of the retrieval process–can occur at different speeds. Frequently used words and recently processed information are recalled faster than occasionally used words that take longer and use more WM resources. Ecke has linked this process to neuroscience, in which uncommon words do not activate the neurons and their associated components as much as commonly used words. Words rarely used reduce the connectivity between neurons and nodes, which results in a failure to retrieve information, i.e., forgetting.

2.5 Chapter Summary and Research Questions

Dewey (1929, 1959) and Skinner (1965) provided sound philosophies for pedagogic practice that have served as a solid foundation on which further theories have developed (Baddeley & Hitch, 1974; Cowan, 1999; Mayer, 2005; Moreno, 2006). Over time, theorists have identified that the traditional paper-based educational systems may not always retain the interest of students, or provide for the needs of a changing society (Li, 2007; Looi et al., 2010; Means et al., 2009). However, learning using M-technology could enhance educational reforms and learning outcome (Bennett, 2012).

The primary purpose of instruction is to build representations of information in LTM, yet it has been ascertained that the capacity for information to be processed is limited (Sweller, 1994, 2010). Intrinsic cognitive load processes core information that has to be understood for learning to take place; it is the intellectual complexity of resources and cannot be changed (Sweller, 2010; Sweller et al., 1998). Reduced extraneous load can increase WM capacity (Wong et al., 2012). In addition, instructional design has been identified as having an influence on learning (Sweller, 2010). According to Mwanza-Simwami (2013), design combines abstract concepts with day-to-day processes, for example, students learning about the heart using resources that have been created with labels situated close to the area they are describing.

The theories of learning described in this chapter provide insight into possible ways to improve teaching practice, while recognising the individual differences of gender, WM, and motivation. Pérez-Sanagustín et al. (2017) called for more research in the areas of humanities and sciences, and so the current study aimed to look at learning across a range of educational domains. Due to the ubiquitous nature of technology (CISCO, 2017), understanding how to utilise M-technology as a tool to improve student learning was an important consideration of this study. Findings could lead to informed methodologies of learning for the future development of M-technology and its use in education.

The research questions were:

- Does mode of presentation (paper, M-technology) affect learning outcome?
- Does cognitive load affect learning outcome?
- Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?

Chapter 3. Methodology

This chapter includes information relating to a methodology rationale. It contextualises the research and elucidates the processes of research design and methods, data collection procedures, and the process of analysis. The purpose of this study was to determine differences in learning when information is presented to students on different media (paper, M-technology), whether learning outcome differs if information is presented in different forms (text, graphics, audio), and to ascertain the effect individual differences have on learning.

The research questions were:

- Does mode of presentation (paper vs. M-technology) affect learning outcome?
- Does cognitive load affect learning outcome?
- Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?

3.1 Paradigm Rationale

According to Raadschelders (2011), ontology speculates about the nature of reality and engenders theories about "what can be known (epistemology), how knowledge can be produced (methodology), and what research practices can be employed (methods)" (p. 920). This is in line with Crotty (1998) who identified an ontological research paradigm consisting of epistemology, theoretical perspective, methodology, and methods. Scotland (2012) recognised the importance of identifying an ontological approach in order to understand the perspective of the research. Two key ontological approaches, i.e., constructivism and critical realism, were considered for the current study and are discussed below.

Constructivism is an approach that has evolved over time. Piaget (1967) identified personal constructivism where knowing is an adaptive activity, i.e., through actions children construct their own development. Later, Vygtosky (1978) acknowledged social constructivism as a mental function that is embedded in a sociocultural setting mediated by other people through the use of language, i.e., a zone of proximal development. von Glasersfeld (1989) recognised that an individual's knowledge of the world is bound to personal experiences. It is only by checking our understandings and perspectives with others that we develop a sense of the viability of ideas (critical thinking). However, viability (unlike truth) is relative to a context, but not limited to the concrete or material–we construct within conceptual contexts (von Glasersfeld, 1995). Learning comes from conflict between existing concepts and new knowledge–it is possible to hold original intuitive views, e.g., people are trustworthy, simultaneously with newly constructed formal concepts, e.g., not everyone is trustworthy. Understanding requires the learner to actively engage in meaning-making. Knowledge is built up from the movements of attention in the dynamic construction of images we create in our minds, and is therefore subjective.

Constructivism has been identified as a reality that can only be known in our experiential world, an approach that is subjective in nature (Von Glasersfeld, 1995). An opposing perspective, i.e., critical realism, is where reality is a universal truth and is not subjective but objective in nature. As the aim of this research was to identify the best conditions for learning, i.e., to identify a reality, an objective approach was called for. Instances of an objective learning approach are: (a) cognisance–when a person is made more aware of specific aspects of the environment and can identify a truth; and (b) habituation–when a person is less responsive to the environment as elements within the environment have become an internalised truth, e.g., prior knowledge (Olson & Hergenhahn, 2016). Therefore, and in line with the very nature of ontology, the underlying framework of the current study was based on the paradigm of critical realism, a philosophical approach in which the world exists and is separate to our knowledge of it, in which truth can be discovered in an objective way through research (Sayer, 1992; Speed, 1991). Critical realism understands that access to a reality is determined by our perceptions. However, a causal effect infers existence irrespective of perception. Critical realism acknowledges an epistemic relativity, i.e., knowledge that occurs within a particular society, in which all points of view are valid and different forms of knowledge–such as physical, e.g., atoms, social (family), and conceptual (ideas) –are embraced. As such, the object of research may have different characteristics, and it is possible that a mixed-method research strategy may be necessary (Mingers, Mutch, & Willcocks, 2013).

Critical realism methodology is abductive in nature, in which empirical research is motivated by theory and hypothetical mechanisms are applied to unexplained phenomenon to provide an explanation for an effect (Meyer & Ward, 2014; Mingers et al., 2013). This demonstrates movement from an empirical domain to a real domain–moving from events to the causes of those events and examining the interaction. There may be multiple interactions and, ultimately, the only way to determine a truth is through research that rejects some of the accounts and supports others. According to Mingers et al., critical realism methodology should “describe the events of interest; retroduce explanatory mechanisms; eliminate false hypotheses; [and] identify the correct mechanisms” (p. 3). As the structure could encompass material, social, and cognitive forms, a variety of research methods is required.

Wynn and Williams (2012) advocated that the focus of causality within critical realism is recognising that human subjectivity exists within certain situations alongside structures, i.e., related objects, e.g., an organisation that direct people’s actions within their environment. The relationship between objects in a structure is unique, and creates events within the structure that are individual to the structure itself. This is beneficial, as a more detailed reason can identify causes of events from the perspective of those within the environment, as well as objects that interact to affect an outcome. Critical realism, therefore, affords a multi-level methodology in which analysing cause and effect states how and why an event occurred. This process makes it more difficult to generalise findings as each circumstance is unique (Zachariadis, Scott, & Barrett, 2013). Wynn and Williams identified that a socio-technical situation includes: individuals; rules and practices; technological tools; and language and culture. These structures interact to generate events. For information systems, empirical research provides causal explanations by identifying the levels at which the tools and devices produce events.

Mir and Watson (2001) argued that a good theory must contain strong assumptions, demanding methodology, and be cautious about the generalisability of findings. Critical realism meets these criteria. In 1989, Bhaskar (1989/2013) identified the core of realist philosophy, in which “perception gives us access to things and experimental activity access to structures that exist independently of us” (p. 9). Critical realism comprises of mechanisms, events, and experiences. Mechanisms that are the basis of causal laws, encompass nature as well as patterns of events that occur apart from experience. Structures are autonomous to events and help to make sense of the investigation of events. When a new layer of reality has been revealed, science tries to make sense of it through research.

According to critical realism, and within the current study, a layer identifying the escalating use of technology in schools was acknowledged and investigated. The factors of mode of presentation and cognitive load were viewed as artefacts that enable an individual to achieve a state of consciousness, i.e., learning. The students’ socio-cultural environment was one of being educated in an international school, all with similar learning experiences. This process supported Healy and Perry (2000) who determined that ontology is the reality that researchers investigate, and epistemology is the relationship between that reality and the researcher. The critical realism paradigm was useful to the current study, as an objective scientific approach was necessary to ascertain a truth to determine the best conditions for recall using different modes of presentation (paper, M-technology), as well as exploring the most effective cognitive load (text-only; text & graphics; graphics & audio; text, graphics & audio). An existing layer of reality needed to be examined to determine a methodology that would enhance learning.

A pedagogy that supports critical realism, and which is relevant to the current study, is the activity theory (Nardi, 1996). Activity theory focuses on understanding consciousness (attention, memory, learning) and activities in the real world. According to Kaptelinin and Nardi (2006), activity theory is "an individual's interactions with people and artefacts in the context of everyday practical activity" (p. 8). Activity theory identifies external influences that support learning. It is not associated with any one philosophy, but over time has been linked to: classical German philosophy; Soviet cultural-historical psychology (Vygotsky); and Dewey's pragmatism. As such, it is multidisciplinary and is connected between disciplines by the central concept of activity (Nardi, 1996). Activity is dynamic as an individual changes, and should, therefore, be observed over a longer period of time. It has been recognised that an individual's development is grounded in a socio-cultural environment (Cote, Kwak, Putnick, Chung, & Bornstein, 2015; LeVine, 2017). However, in activity theory the individual is not reduced to society or culture. In fact, the procedures of internalisation-externalisation allow individuals to alter culture through activity (Kaptelinin & Nardi, 2006). In summary, the current study was based on critical realism and activity theory, which sets out to determine a truth.

3.2 Methodology

Methodology is fundamental to ontology and epistemology, because the suitability of research centres on assumptions about the type of causal relations they are meant to ascertain. There should be congruency between methodology and ontology for greater validity (Hall, 2003). Crookes (2013) acknowledged that social science research has been recognised as being inter-disciplinary, requiring different methods of research that encompass both objectivity and interpretivism, i.e., a “plurality of research methods” (p. 3) that discover how knowledge is attained. The current study set out to discover truth in an objective way, and, in a scientific attempt to access some of these causal relationships, a quantitative dominant with a pure quantitative methods approach was selected (Johnson, Onwuegbuzie, & Turner, 2007). The quantitative method allowed interactions and effects to be examined through statistical analysis (see Figure 3.1).

Figure 3.1. Research Paradigms Identified by Johnson, Onwuegbuzie, and Turner (2007)

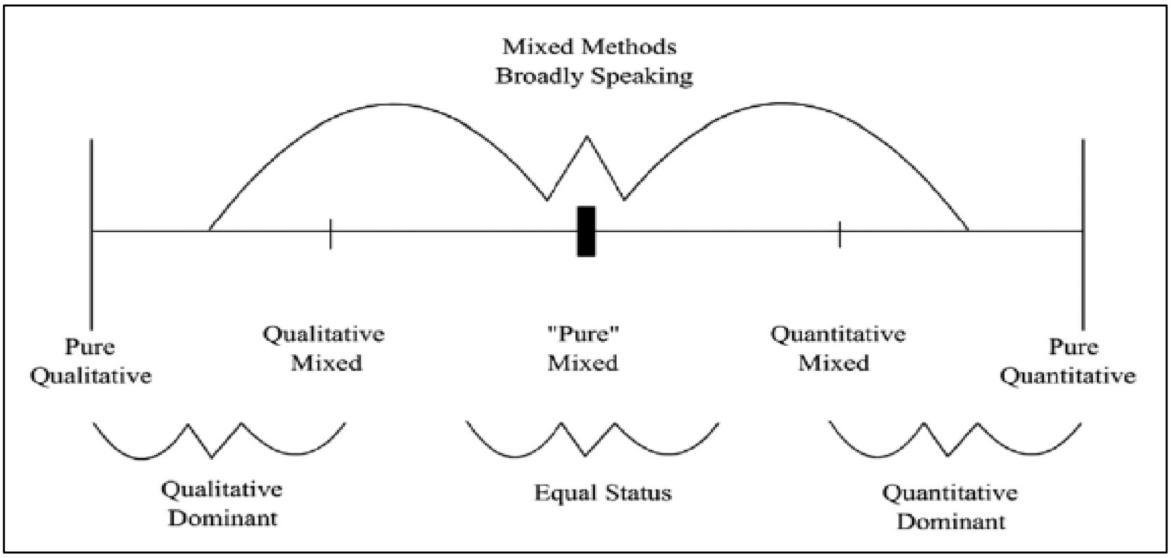


Figure 3.1. Research Paradigms from Johnson, Onwuegbuzie, and Turner (2007)

3.3 Sample Selection

Participants consisted of students from a high achieving multinational school in Qatar, in the Middle East. Over 50 nationalities were represented within the school. With maximum class sizes of 24, the school has an excellent reputation and delivers the British curriculum to mixed gender classes in English. Sixty-seven percent of students’ first language is English, but all speak and understand English to a high level. Even though there are many able students, there is a mixed-ability element as the school is non-selective and has a learning support department. The school is a one-to-one mobile-device school, in which every child and teacher has an iPad®. All KS3 students took part in the research, but eight students' data were not included in the final analysis due to lack of parent consent (N = 5) and student consent (N = 3).

Participants consisted of an opportunity (convenience) sample, comprising of 346 students in KS3 from one school. Student were aged 11-14 years (*M* = 13 years 4 months, *SD* = 0.95), including 167 boys and 179 girls. Even though the selected students represented an age where a great amount of cognitive development occurs (Mirza & Hussain, 2018; Shaffer & Kipp, 2013), they provided a larger sample size than from any other key stage. A large sample size was necessary to ensure there were enough participants in each of the testing conditions, e.g., four cognitive load conditions (text-only; text & graphics; graphics & audio; text, graphics and audio) and two motivation conditions (high, low). Field (2014) identified a sample size of 10-15 participants per variable was adequate. In addition, sampling error could be an issue, i.e., errors that arise in statistical analyses due to smaller sample sizes that do not fully represent of the whole sample. Bryman (2016) identified that the greater the total number of participants taking part in the research, the smaller the sampling error. Bryman also recognised that the heterogeneity of the population, i.e., differences in a population, could help to determine sample sizes, "when it [a population] is relatively homogeneous, such as a population of students… the amount of variation is less and therefore the sample can be smaller" (p. 186).

Opportunity sampling is advantageous in that it allows research to take place with participants who are freely available, thereby taking less time and/or cost to choose participants. It is a quick, easy, and economical way to find participants for research. The current study's participants were a non-randomialised sample, as not every person in the whole population had a equal chance of being selected. A disadvantage of opportunity sampling is that the sample is not representative of the whole population, and results would be valid for that population only. Therefore, generalisation would be an issue (Cohen, Manion, & Morrison, 2017). However, we do know that the sample in the current study represented a multicultural, high achieving international group of students in a British curriculum environment, and so the findings may be generalised to those specific populations. Findings could support/refute research conducted in similar settings or be used as a trigger for future research (Bryman, 2016).

3.4 Tasks, Materials, and Measures

The current study was a quasi-experimental study where the research took place in a natural setting, i.e., classrooms within a secondary school. The independent variables of mode of presentation (paper, M-technology) and cognitive load (text-only; text & graphics; graphics & audio; text, graphics and audio) were manipulated where some participants were randomly placed in the paper sample groups and others in the M-technology sample group (Cohen, et al., 2017). Sample groups were not matched in any way except that there was a Year 7, Year 8, and Year 9 class in each condition.

Prior research has pursued educational topics, such as the technical process of hydraulic brakes (Mayer et al., 2007) and geographical weather maps (Lowe, 2003). Similarly, the current study chose to explore educational topics across different subject disciplines, to determine whether outcome is similar across subjects. Tasks consisted of: (a) science–the heart, which included technical language that described the process of blood circulation; (b) geography–a map that assessed spatial awareness by including a story line to aid recall; and (c) English–a rhyming poem (Greek fable), that was largely text-based. In the past, research has tested memory using word pairs or small amounts of data (Fenn, Nusbaum, & Margoliash, 2003; Gathercole, Hitch, Service & Martin, 1997), but more recent research has started to examine learning outcome using longer texts (Mangen et al., 2013). In line with present trends, the current study tested learning on an extended volume of data, which is more representative of the type and amount of material participants process at school (see Table 3.1., Table 3.2., Table 3.3.).

Falloon (2013) argued that appropriate design parameters should emulate similar learning structures as those provided by a teacher. As such, design and justification for choice of design for the current study is discussed.

3.4.1 Design

The research design can affect the “validity of any conclusions that might be drawn from the research” (Ary et al., 2018, p. 271). An experimental research design ensures at least one variable can be manipulated while other extraneous variables are controlled, establishing a cause-and-effect relationship between variables (Gravetter & Forzano, 2018). However, in a natural setting this is difficult to accomplish and a quasi-experiment would need to be conducted, where confounding variables cannot always be controlled, e.g., elements of the environment such as interruptions during the testing period (fire alarm), can threaten the internal validity of the research. The current study was categorised as a quasi-experiment (in which it is difficult to determine a cause-and-effect relationship between variables). Even though not all variables could be controlled, as many known variables as possible were taken into account. For example, instructions were consistent between sample groups, testing periods were timetabled when students were not out on school trips, and the data collection for the second testing time occurred for all students on the same day at the same time. These procedures ensured internal validity was considered.

Within the current study IVs, i.e., the predictor variable, were mode of presentation (paper, M-technology), cognitive load (text-only; text & graphics; graphics & audio; text, graphics & audio) and individual differences (gender, WM, motivation). Individual differences were considered extraneous variables. The IV was used to answer the research question, "Does mode of presentation (paper, M-technology) affect learning outcome?", "Does cognitive load affect learning outcome." and "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?" Gender, WM, and motivation were identified as additional extraneous IVs that may have influenced findings, and were included in the study to determine the degree to which they moderated the key experimental variables of mode of presentation and cognitive load. Participants were randomly assigned to a paper or M-technology condition by tutor groups, one from each year group (Years 7-9). The DV, i.e., recall, identified the amount of information participants could recall after 40 minutes of being presented with the learning material and again four to six weeks later. The IVs were nominal and contained two categories while the DV data were continuous, i.e., incremental.

As the current study was a quasi-experiment, it was difficult to ensure the IVs were the only variable to affect the dependent variable. Extraneous variables, such as time of day, emotions of the participants, immersion with the content of the material, interruptions, etc. were more difficult to control than if the research were purely experimental.

3.4.2 Material overview

Within the current study, design criteria were considered and extraneous processing was reduced by: formatting text, where possible, to be a sans serif font; highlighting key words by enlarging the font and using bold and underline; ensuring text was left aligned and in block form; and situating labels close to items they were identifying (Abeysekera & Dawson, 2015; Moreno, 2007; Norman, 2013; Schlag & Ploetzner, 2011). This was to reduce the amount of extraneous cognitive processing, and allow the true measurement of cognitive load to be ascertained to be able to answer the research question, "Does cognitive load affect learning outcome?"

When considering which applications (apps) to use, Falloon (2013) identified the importance of considering the factors of cost, rating, and meeting the requirements of the research. The current study selected apps with an online rating of four out of five stars, that were cost-free, and were, or had the potential to be, multi-modal (text, graphics, audio), so they could be adapted for different levels of cognitive load. The Visual Body® app, which consisted of a video and audio of the function of the heart, was identified for the heart curriculum topic. The TabTale Fox and Crow® app, a rhyming poem that included text, graphics and audio, was selected for the poem curriculum topic. An additional app (WebDAVNav®) was used to enable participants to save work from their iPad® to the school shared drive. This app was used to facilitate the submission of work. In addition to giving participants verbal and written instructions, WebDAVNav® was relatively intuitive to use, an important component identified by Lynch and Redpath (2014). No appropriate existing app was found for the map task, so resources were created for this task.

Control and autonomy was considered for the current research, and, in accordance with existing research, all the testing material included design features that were considered beneficial for learning (Abeysekera & Dawson, 2015; Falloon, 2013; Lynch & Redpath, 2014). This included a button in each of the topics (heart, map, poem), from which the resource could be played, paused, reversed, and forwarded. In the map and heart tasks, signalling took the form of formatted text to identify key words (enlarged and/or made bold). In the poem task, each line of onscreen text was highlighted simultaneously with the audio. In the heart task, text appeared on screen at the same time as the voiceover, and the animation zoomed in to relevant sections being discussed (Norman, 2013).

Even though graphic animations may have included extraneous information, such as colour and movement (Ayres & Paas, 2007), in the current study colour was retained, as it is the norm for students in an educational setting–to remove it would create a false environment and reduce the validity of the findings. It has been argued that dynamic graphics (animation) can be less effective than static graphics, because they are viewed for a much shorter time and use WM resources (Ayres & Paas, 2007; Yung & Paas, 2015a). This parameter was considered in the current research, and very little animation was used. There was no animation in the map and poem tasks, and minimal animation in the heart task.

It has been established that encoding is more difficult for high level vocabulary and, therefore, more difficult to retrieve (Gorin & Embretson, 2006; Graves, 1986). Within the current research there was some unfamiliar vocabulary, e.g., pericardium in the heart task and Haight street in the map task. The current research endeavoured to select material composed of high intrinsic cognitive load and low germane load.

3.4.2.1 Heart Task

3.4.2.1.1 Learning phase materials

An animated audio-video of the function of the heart was downloaded from the free access section of the VisibleBody Atlas® app. The audio was transcribed for both paper and M-technology resources (see Table 3.1.). A similar graphic of a heart to the one in the animated video was found for the paper condition. This was to ensure consistency between paper and M-technology resources so that the comparison between paper and M-technology learning outcome was accurate when answering the research question, “Does mode of presentation (paper, M-technology) affect learning outcome?”

Table 3.1.

Material for the Heart Task Identifying the Research Design with Data for Mode of Presentation (Paper, M-technology) and Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, graphics & Audio) Presented Separately

		Mode of Presentation				Cognitive Load			
	Words (Time)	A4 Paper		M-technology iPad		M-technology (iPad, Keynote® Presentation)			
	N	T	TG	T	TG	T	TG	GA	TGA
Design	184 (68 s)	Black text on white paper	Coloured graphic, black text, on white paper	Black text on white screen	Coloured graphic, black text, on white screen	Black text on white screen	Black text, animated graphics	Animated graphics, American voice-over	Black text, animated graphics American female voice-over

Note. T = text-only, TG = text & graphics, GA = graphics & audio, TGA = text, graphics & audio. Fonts in the testing material consisted of sans serif font. See Appendices B-D.

In the M-technology cognitive load research, design choices included: (a) control, in which participants were able to manipulate the speed by forwarding/rewinding the presentation, and scrolling up/down to read text; and (b) signalling, where the text and audio were synchronised with the animation. These features are part of the cognitive load processing, and help to provide an answer to the research question, "Does cognitive load affect learning outcome?"

3.4.2.1.2 Test phase materials

There was a prior knowledge assessment for this topic. Participants were provided with a graphic of the heart that could be labelled, and given the opportunity to write down what they knew about the heart, a total of 9 marks were awarded. Papers were collected and students then given the testing material.

The immediate assessment was composed of two parts: Section A required the recall of facts and Section B asked participants to apply their knowledge (see Appendix E). A total of 14 marks were awarded (9 marks for factual knowledge and 5 marks for applying the facts to a scenario). In Section A, the graphic of the heart had seven labels. As some of the labels were similar, one mark was awarded for achieving both. This occurred for left/right atrium, and left / right ventricle. Additional marks were awarded for: oxygenated/deoxygenated blood (an additional mark was given if labelling was on the correct side of the heart); pericardium (sac, wall, layer, or lining, was accepted); and any additional information about the heart (the shape, the fact that it was an organ, and acted as a double pump). Changes were made to two questions from the time of immediate recall to delayed recall. This was to ensure learning had occurred and participants did not just remember the answer they had initially given.

Immediate recall:

- Aseel buys a 500g tub of butter. How much more is this than an adult heart? Correct answer, 200g.

Delayed recall:

- Omar buys a 1kg tub of butter. How much more is this than an adult heart? Correct answer, 700g.

The assessments were triple marked to verify the scores. The data were entered into a spreadsheet and the subtotals, totals, and percentages were calculated. The data were then entered into SPSS® software for further analysis.

3.4.2.2 Map Task

3.4.2.2.1 Learning phase materials

The map resource was created by combining clipart graphics to construct a single graphic scenario (map, church, greenhouse, castle, school, dog, park, zoo, burnt tree, supermarket, hospital, dinosaur, and children). The presentation format was composed of a static graphic within a single frame. Story lines were created that explored different areas of the map and allowed participants to trace routes (see Appendices F-H). Text and graphics were consistent between paper and M-technology resources so that the comparison between paper and M-technology learning outcome was accurate when answering the research question, "Does mode of presentation (paper, M-technology) affect learning outcome?"

Table 3.2.

Material for the Map Task Identifying the Research Design with Data for Mode of Presentation (Paper, M-technology) and Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, graphics & Audio) Presented Separately

		Mode of Presentation				Cognitive Load			
	Words (Time)	A4 Paper		M-technology iPad		M-technology (iPad, Keynote® Presentation)			
	N	T	TG	T	TG	T	TG	GA	TGA
Design	400 (134 s)	Black text on white paper	Coloured graphic, black text on white paper	Black text on white screen	Coloured graphic, black text on white screen	Black text on white screen	Black text, animated graphics	Animated graphics, American voice-over. Low instrumental music	Black text, animated graphics South African female voice-over. Low instrumental music

Note. T = text-only, TG = text & graphics, GA = graphics & audio, TGA = text, graphics & audio. Fonts in the testing material consisted of sans serif font. See Appendices F-H.

Design choices included: (a) control, in which participants were able to manipulate the speed by forwarding and rewinding the presentation; and (b) signalling, in which the font size of a key street, Haight St., was enlarged and made bold as it was the starting point for participants when reading the map, and the four main points of a compass (N, S, W, E) were included outside the frame of the main graphics to orient participants. These features are part of the cognitive load processing, and help to provide an answer to the research question, “Does cognitive load affect learning outcome?”

As the geography map curriculum topic was a novel task involving an invented location, prior knowledge was not considered relevant to students’ performance. Therefore, no prior knowledge existed for this task.

3.4.2.2.2 Test phase materials

The map assessment was composed of two parts: Section A required the recall of facts; and Section B asked participants to apply their knowledge (see Appendix H). A total of 14 marks were awarded (12 marks for factual knowledge, and 2 marks for applied knowledge). There was no prior-knowledge assessment. Changes were made to two questions from the time of immediate recall to delayed recall.

Immediate recall:

- How would Sarah and John walk to the supermarket from their house? Describe their route. Correct answer, walk along Haight St.

Delayed recall:

- If Sarah and John walked north of the park, what would they see ahead? Correct answer, supermarket.

The assessments were triple marked to verify the scores. The data were entered into a spreadsheet and the subtotals, totals, and percentages were calculated. The data were then entered into SPSS® software for further analysis.

3.4.2.3 Poem Task

3.4.2.3.1 Learning phase materials

Aesop’s Fable app, the Fox and Crow, was installed on iPads®. The presentation format was static graphics and a rhyming poem, with four to six lines of text per frame. (see Table 3.3.). Text and graphics were consistent between paper and M-technology resources so that the comparison between paper and M-technology learning outcome was accurate when answering the research question, “Does mode of presentation (paper, M-technology) affect learning outcome?”

Table 3.3.

Material for the Poem Task Identifying the Research Design with Data for Mode of Presentation (Paper, M-technology) and Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, graphics & Audio) Presented Separately

	Words (Time)	Mode of Presentation				Cognitive Load			
		A4 Paper		M-technology iPad		M-technology (iPad, Keynote® Presentation)			
	N	T	TG	T	TG	T	TG	GA	TGA
Design	516 (255 s)	Black text on white paper	Coloured graphics, textbox with black text on white background set in the graphic at the top, on white paper	Black text on white screen	Coloured graphics, textbox with black text, on white background set in the graphic at the top, on white screen	Black text on white screen	Brown text Each line highlighted for a short duration, static colour graphics	Static graphics, American voice-over, background music	Brown text, Each line highlighted for a short duration, static graphics American voice-over, background music

Note. T = text-only, TG = text & graphics, GA = graphics & audio, TGA = text, graphics & audio. Fonts in the testing material consisted of sans serif font. See Appendices I-L.

Design choices included: (a) control, in which each frame had arrows that allowed the reader to move forward and backward; and (b) signalling, where individual lines of the poem were highlighted when the audio sounded. The electronic app's graphic frames (see Appendix J) were print-screened and used for the paper text & graphics condition (see Appendix K). These features are part of the cognitive load processing, and help to provide an answer to the research question, "Does cognitive load affect learning outcome?"

3.4.2.3.2 Test phase materials

There was a prior knowledge assessment for this topic for which a total of 6 marks were awarded. Participants were provided with general questions from the Fox and Crow story and were asked to answer as many questions as they could. The papers were collected, and then the testing material was handed out for students to learn from. The poem assessment was composed of two parts: Section A required the recall of facts; and Section B asked participants to apply their knowledge (see Appendix L). A total of 13 marks were awarded (7 marks for factual knowledge, and 6 marks for applied knowledge). The assessments were triple marked to verify the scores. The data were entered into a spreadsheet and the subtotals, totals, and percentages were calculated. The data were then entered into SPSS® software for further analysis.

3.4.3 Motivation questionnaire

3.4.1.1 Testing phase materials

The current research endeavoured to examine the individual difference of motivation to determine whether motivation had an impact on recall. Participants completed an adapted version of Pintrich et al.'s (1991) motivation questionnaire. The number of items were reduced to 18 motivational items to reduce participant fatigue (see Appendix M). Items were scored on a 7-point Likert-type scale, from 1 (not true) to 7 (very true). Motivational categories included four questions each on goal orientation, task value, effort, and self-efficacy. Two questions were added to the category of test anxiety, but were not included in the data analysis as they did not pertain directly to motivation. The questionnaire was used to help answer the research question relating to whether individual differences in motivation moderate the effect of mode of presentation and cognitive load on learning outcome.

According to Barnette (2000), participants have a tendency to agree with survey statements. A way to guard against acquiescent behaviour is to reverse the wording, in which questions are worded opposite in meaning to the majority of the questions. However, Van Sonderen, Sanderman, and Coyne (2013) and Suárez-Alvarez, Pedrosa, Fernández (2018) believed reverse-worded questions only served to confuse participants. Therefore, in the current research, three questions were reverse-worded where participants had to carefully consider their answers. Before analysis, scores were reversed to allow for the consistency of scale values for all responses.

The questionnaire adhered to the parameters of good design, i.e., consistent spacing, contrasting text against background, no colour, numbering, scaffolding with signals to remind one of scale values (not true to very true) for each question. In addition, the scale of 1-7 was situated below each question demonstrating consistency in layout (Abeysekera & Dawson, 2015; Moreno, 2007; Norman, 2013; Schlag & Ploetzner, 2011).

3.4.1.2 Post-testing phase calculations

To determine levels of motivation, data were inputted into a spreadsheet and the mean of all questions for each participant was calculated, after which the median of all the raw scores was determined. All scores above the median were classified as high motivation and assigned a value of 1, and all scores equal to or lower than the median were classified as low motivation, and assigned a value of 0. Data were then inputted into SPSS® for further analysis.

3.4.4 Working memory assessment

3.4.4.1 Testing phase materials

Lucid Recall®, an electronic assessment of WM for students 7-16 years-of-age, was used to assess the participants' working memory. The Alloway Working Memory Assessment (AWMA), was considered, but not used as it was designed for individual administration that is extremely time consuming, and the current research was a large scale study. Lucid Recall® was selected due to ease of administration, i.e., it was fully automated and up to 20 participants could be assessed at the same time. Working memory was considered as a key source of individual differences that might moderate the effects of mode of presentation and cognitive load on recall.

The Lucid Recall® assessment consisted of three separate tasks for the different WM structures: The phonological loop was assessed with a word recall task; the visuo-spatial sketchpad was assessed with a pattern recall task; and the central-executive function was assessed with a counting recall task (see Figure 3.2). Each task took 10-15 minutes depending on the participants' WM capacity; the assessment was designed to increase the complexity based on previous correct answers. Tasks began with an audio demonstration to ensure understanding. Headphones were provided, and participants completed the assessment in silence.

Figure 3.2. Lucid® Working Memory Tasks

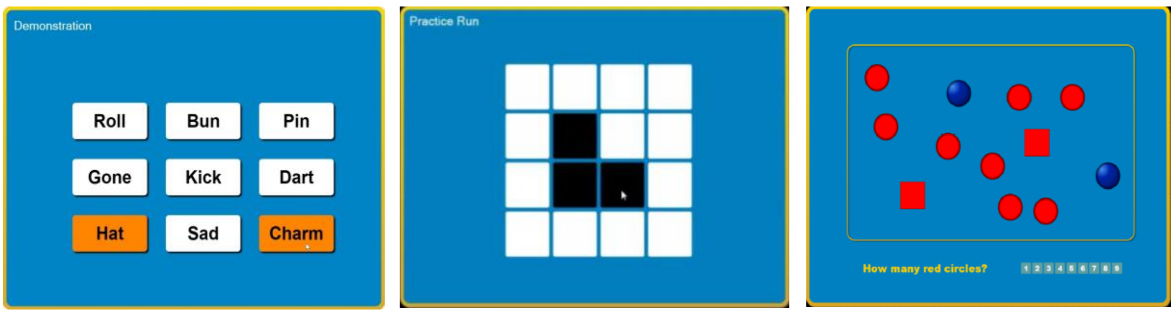


Figure 3.2. Screens showing the word recall task (phonological loop), pattern recall task (visuo-spatial sketchpad), and counting recall task (central executive).

The software was installed on the school's network system and the participant names were imported from a spreadsheet. Participation was voluntary. Data collection took place at opportune times after the delayed recall assessment, up until the end of the school year.

3.4.4.2 Post-testing phase calculations

Lucid Recall® is an automated self-marking application. To determine participants' levels of WM, data were exported to a spreadsheet and the median of the composite raw scores was calculated and used as a measurement of WM. All scores above the median were classified as high WM and assigned a value of 1, and all scores equal to or below than the median were classified as low WM, and assigned a value of 0. Data were then inputted into SPSS® for further analysis.

3.4.5 General Information

3.4.5.1 Assessment and marking

In the current study, participants completed an information form–one before the testing period and one after the testing period. The initial form gained the participants' consent, as well as identified the participants' home language, how they felt about taking part, and how they felt they learnt best (see Appendix N). The final questionnaire included a question that identified whether students revisited the questionnaire between testing periods, and, if so, how many times they read through the material (see Appendix O).

All data were inputted into a spreadsheet and analysed. The question "What is your home language?" provided information of the number of different cultures represented within the sample, as well as the number of participants for whom English was a second language, which could have impacted on recall.

3.5 Procedure

3.5.1 Timeline

Organisation and planning are important components for success. Thomas (2017) found that more time spent planning resulted in greater performance. As such, a timeline was planned for the current study (Figure 3.3.). However, a flexible approach was necessary with the option for rescheduling procedures, due to any unforeseen circumstances that may have occurred. The timetable for data collection was scheduled around the school calendar, so activities such as school trips disrupted the research as little as possible.

Figure 3.3. The Current Study's Research Timeline

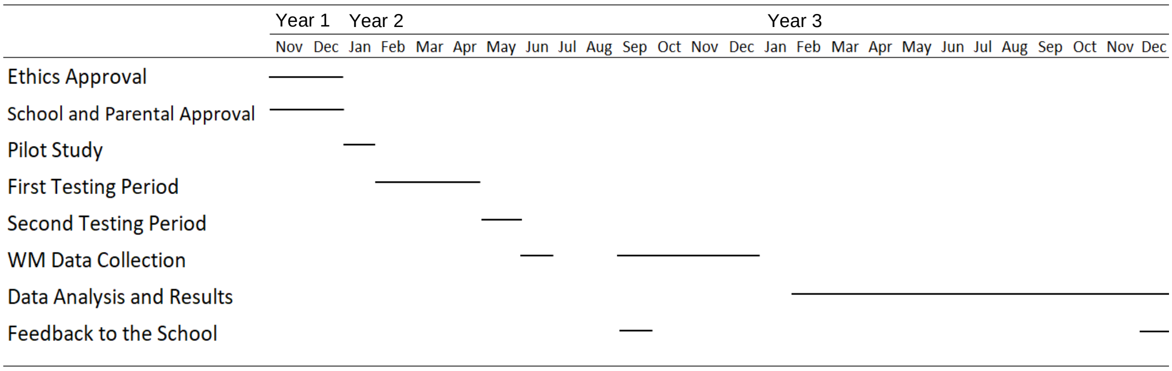


Figure 3.3. The current study's research timeline, identifying key points in the research process.

3.5.2 Pilot study

In line with the human computer design framework that supports observation, conceptualisation, prototyping, and testing (Norman, 2013), a limited pilot study was conducted with one Year 10 class (N = 12) to identify whether any general design corrections needed to be made, and to determine whether the researcher and participant perceptions of the research matched (Hammersley & Traianou, 2012). A Year 10 class was selected as it did not draw on participants from the sample group who were needed to maximise numbers for the research, and they were closest in age to the sample group.

The Year 10 participants identified areas for improvement in the design of the research material (Walliman & Buckler, 2008). For the heart task, participants had written a fair amount in the prior knowledge section, and were reluctant to rewrite the information again in the immediate recall testing time. As a result, the questionnaire for the immediate recall testing time was changed to “Write as much as you know about how the heart works that is different than you wrote before.” Also, participants did not like the word “assessment” at the top of the page as they felt it resembled regular school tests, for which they would be negatively judged if they performed poorly–they suggested using the word “quiz” instead. A further factor to consider was identified, i.e., potential issues with the use of technology. In one instance, an iPad® was not fully charged, and in another the device had been dropped and the cracked screen had rendered it useless. It was, therefore, worthwhile to have some additional backup devices. After completing the pilot study, participants were relaxed and acknowledged they were pleased to be asked their opinion. In terms of planning for the actual data collection, the pilot study was invaluable and, as a result, yielded more reliable data.

3.5.3 Data collection

A quasi-experimental methodology was the primary source of empirical data collection in the current study as the environment could not be fully controlled. Performance on different tasks was measured in a natural environment, i.e., the classroom. However, even though the research did not take place in a laboratory, where most variables can be fully controlled, there were enough participants taking part to negate the impact of extraneous variables and allow for general trends to appear (Cotton, 1995).

As the current research was a situated study, the results are likely to be more generalisable than in a laboratory. By comparing findings to existing and future research, interactions between events can be examined and help to identify a truth, thereby identifying a cause and effect. Situations may yield similar results, after which generalisations may be made.

Ethical approval was gained from the university, and permission for the research to take place given by the school (see Appendix A). Letters were then sent out to parents. On the recommendation of the school principal, and due to the historically poor response of parents, the opt-out method was used. Prior research has shown a greater response for opting out than opting in (Hunt, Shlomo, & Addington-Hall, 2013; Vellinga, Cormican, Hanahoe, Bennett, & Murphy, 2011). Consent was gained from participants in the initial questionnaire. All KS3 participants, Years 7-9 inclusive, took part in the research data collection, but the data of parents and participants who had opted out was disregarded.

There were six classes in each of the three year groups, with a maximum of 24 students in each class. Classes represented different houses that were identified by a character (D, G, P, R, S, and U). As sampling groups were set by class, the number of males and females per group was pre-determined, and were, therefore, not always equal. Additional variance in number of participants per condition was due to absences and technology issues. Therefore, at times group sizes were unequal, or very small. Classes were randomly allocated to testing conditions for mode of presentation (paper, M-technology) and cognitive load (text-only; text & graphics; graphics & audio; text, graphics & audio). They were not matched in any way (see Table 3.1.). Counter-balancing was implemented by allocating one class of each year group to the sample for each testing condition, so there were a similar number of Year 7-9 participants in every group. To avoid a practice effect, sample groups did not repeat the same paper or M-technology condition. The cognitive load research included only the M-technology sample group data, to test Sweller's (2008) CLT and Mayer's (2005) CTML.

The first data collection testing time, immediate recall, took place between February and March of the data collection year, and the second testing time, delayed recall, four to six weeks later in the first week of May.

Table 3.4.

The Allocation of Sample Groups and Conditions for Each Task, with Data for Paper Presented Separately to M-technology, to Investigate an Answer to the Research Questions, "Does Mode of Presentation (Paper, M-technology)Affect Learning Outcome?" and "Does Cognitive Load Affect Learning Outcome?"

Task	**Paper Condition**				**M-technology Condition**							
	Text		**Text, Graphics**		**Text**		**Text, Graphics**		**Graphics, Audio**		**Text, Graphics, Audio**	
	Tutor Grp	**N**	**Tutor Grp**	**N**	**Tutor Grp**	**N**	**Tutor Grp**	**N**	**Tutor Grp**	**N**	**Tutor Grp**	**N**
Heart	7P	56	7R	52	7G	68	7S	35	7D	49	7U	51
	8P		8G		8R		8D		8S		8U	
	9P		9G		9R		9D		9S		9U	
Map	7S	57	7D	34	7U	59	7G	46	7P	47	7R	49
	8G		8P		8U		8R		8D		8S	
	9R				9U/9P		9G		9D		9S	
Poem	7D	43	7S	55	7P	58	7R	56	7U	46	7G	48
	8S		8D		8G		8U		8P		8R	
	9S		9R		9P		9U		9G		9D	

Note. Tutor Grp = Year group (7, 8, 9) and House (D, G, P, R, S, U).

A carousel system ensured that the researcher conducted the majority of the research with all participants, which helped to control extraneous variables, e.g., instructions given, and thereby ensure reliability of the research process. Data collection took place during time-tabled Information Technology (IT) lessons over a six week period. However, to ease the disruption to IT lessons, the mathematics' department offered their support, and one set of data were collected by mathematic teachers on the same day at the same time. Mathematic staff were given set instructions to follow under the guidance of the researcher, and were monitored throughout the data collection period (see Appendix P).

Participants were asked to install the app WebDavNav® on their iPads® in advance of the research. The app enabled them to save work directly onto the school's network drive, which allowed the technology data to be collected efficiently immediately after testing. After each testing condition, participants were asked to delete the resource to prevent them from reviewing the data between testing periods. However, this was difficult to monitor and participants could easily have not deleted the app or download the app again later on. The traditional paper-based method was examined alongside M-technology, and the paper-based material was collected in so participants did not retain a copy.

Data collection took place over two testing times. At the first time of data collection, in the heart and poem tasks, participants were given a five minute prior knowledge quiz, after which they were provided with the material for the testing condition. They were asked to download the required poem app, or open a file that had been e-mailed to them prior to the start of the lesson. As complex vocabulary and lengthy text cannot be learnt with a single repetition (Ebbinghaus, 1923/2013), participants were given five minutes to read/listen (in silence) after being instructed to review the material as many times as possible during these five minutes. Participants then resumed their normal IT/Mathematic lesson. Thirty-five to forty minutes later, participants were given the immediate recall assessment, and allowed up to 10 minutes to complete it, in silence.

The second data collection took place four to six weeks after the initial data collection. All participants were assessed on the same day at the same time, i.e., the start of the day, 07.15. Volunteer staff were given a set of instructions to read out (see Appendix P), and data collection took place in examination conditions. Participants completed an adapted Likert style motivation questionnaire (Pintrich et al., 1991) straight after the assessment. Assessments were marked three times and data entered into SPSS® software for statistical analysis. The research was quantitative dominant (Johnson, Onwuegbuzie, & Turner, 2007) and the data were analysed via inferential statistics to identify levels of significance. After the second data collection point, volunteer participants completed the Lucid Recall® online WM assessment that took place over a two month period. Data were imported into SPSS® for analysis. Working memory scores were used to allocate participants into high and low WM groups, as described in section 3.4.4.2.

3.6 Analysis and Presentation of Results

Halpern (2013) proposed that quantitative experimental research is a way to clarify relationships amongst variables and allow significant interactions and effects to be determined (Johnson et al., 2007). Quantitative data were obtained in the current study and were analysed using SPSS®.

3.6.1 Statistical analysis

Data were inputted into the IBM SPSS Statistics® application. For each condition, exploratory analysis identified whether the data were normally distributed. The mean and standard deviation scores also provided an initial impression of the data.

A one-way ANOVA has one independent variable that splits the sample into two or more groups, and then identifies whether these groups differ significantly from each other (Field, 2014). In the current research, the independent variable for the first research question, “Does mode of presentation (paper, M-technology) affect learning outcome”, was mode of presentation (paper, M-technology), and the independent variable for the second research question, “Does cognitive load affect learning outcome”, was cognitive load (text-only; text & graphics; graphics & audio; text, graphics & audio). Individual differences were independent variables in both data analyses were used to answer the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?”

ANCOVA analysis was conducted in tasks that assessed the prior knowledge of participants (heart, poem). The effect of the covariate factor (prior knowledge) was removed, which allowed the influence of the independent variables on the dependent variables to be determined.

Factorial ANOVA analysis was conducted when there was more than one independent variable to determine if an interaction existed between the two independent variables and the dependent variables, i.e., when introducing the variable of individual difference (gender, WM, motivation). This determined whether two or more groups–mode of presentation/cognitive load and individual difference–differed from each other significantly in characteristics. However, for the results of ANOVAs to be reliable, the following assumptions needed to have been met:

- The dependent variable should be a continuous interval, i.e., when distances between values are the same, e.g. percentage, from one to two is the same as from eight to nine. The data met this requirement.
- The independent variables should consist of two or more categorical, independent groups. The data met this requirement.
- Independence of observations, in which groups comprise of different participants. The data met this requirement.
- There should be no significant outliers. The data did not always meet this requirement. In these instances, and to reduce bias, the extreme outliers were Winsorised, in which cases data were replaced with the next highest score that was not an outlier (Field, 2014).
- The dependent variable should be normally distributed. Descriptive analysis was conducted, and in most cases this assumption was met. In instances where it did not, non-parametric analysis was conducted. However, it is acceptable for an approximation of normal distribution, as ANOVAs are robust to violations of normality (Field, 2014).
- Groups need to meet homogeneity of variances (when differences are consistent within each group), is for each independent variable combination. Levene's test for homogeneity of variances was conducted, and when results did not meet this assumption, were reported. In these instances, non-parametric analysis was performed.

ANOVAs identified whether the analyses of variance were significant, but to determine which groups' means were different when there were three or more categories, and when no prior knowledge was tested, Tukey's post-hoc HSD test was used to do a pairwise comparison on the data (Field, 2014). Tukey, an adjusted means analysis, took into account the potential issue of Type 1 errors (Type 1 errors occur when the null hypothesis is rejected, i.e., when results of the analysis shows that a significant difference exists when, if fact, it does not. When conducting multiple analyses, there is a chance of a Type 1 error occurring. Tukey and Bonferroni adjustments takes this into account and ensure the results are accurate). When prior knowledge was assessed, Bonferroni's adjusted means test was used, with the prior knowledge data entered as a covariate.

Bonferroni investigated the pairwise differences between cognitive load conditions, while taking into account Type 1 errors.

For non-normally distributed data, the Wilcoxon signed-rank and Mann-Whitney tests, the equivalent of the parametric independent *t*-test, examined the differences between related samples. The Wilcoxon signed-rank test analysed the two-way interaction between each individual differences and time, and the Mann-Whitney test analysed the three-way interaction between mode of presentation/cognitive load, individual difference, and time. The Kruskal-Wallis test, equivalent of analysis of variance, examined differences in the mean ranked positions of scores across all cognitive load groups, while making pairwise comparisons.

3.7 Reliability and Validity

Reliability and validity are important components of research. Heale and Twycross (2015) determined that validity is the extent to which a concept is accurately measured, while reliability relates to the consistency of a measure. Ary, Jacobs, Irvine, and Walker (2018) acknowledged the difficulty in identifying validity: reliability can be determined by examining the test data using analysis and calculations, but validity, of which there are many types, e.g., internal, external, content-, criterion-, and construct-related, etc., requires judgement. Ary et al. identified the relationship between reliability and validity, “If a measure is to yield valid score-based interpretations, it must first be reliable” (p. 120). Reliability and validity, as it applied to the current study, is explained in the following sections.

3.7.1 Reliability

Ary et al. (2018) viewed reliability as being “concerned with the effect of error on the consistency of scores” (p. 102), which may result in an increase or decrease in an individual participant’s score. Ary et al. acknowledged two different types of error. The first is random error of measurement, which includes: (a) factors that can affect a participant’s score in each testing period, e.g., variations in participants’ motivation, tiredness, test apprehension etc. It could also include the malfunctioning of resources, such as the iPad shutting down in the testing period due to low battery power; (b) standardised procedures, e.g., ambiguous instructions and inconsistent scoring. Testing conditions could also affect the reliability of data, such as time of day, distractions, heat, etc. which may be felt differently from child to child; and (c) the test itself. If the assessment for recall is too short, scores may not be reliable as a "small sample results in an unstable score" (Ary et al., p. 104). In addition, results need to be constant when measuring recall between two periods of time. The second type of error is a systematic error of measurement, that would increase or decrease all scores within the sample group, e.g., construction noise outside the testing room that would disturb all the students.

Inter-rater reliability is a way to ensure scores from assessments are reliable; scores would not be reliable if the markers were not consistent or accurate. Gwet (2008) identified that inter-rater reliability "quantifies the closeness of scores assigned by a pool of raters to the same study participants. The closer the scores, the higher the reliability of the data collection method." (p. 29). Hallgren (2012) acknowledged the difficulty in grading accurately and consistently between raters, and that coding requires a considerable amount of training and practise to ensure accuracy. Gwet (2014) identified that the process of rating may require some form of judgement, e.g., such as applying a rating scale to performance, and that when two or more raters agree, they can be used interchangeably with the knowledge that the outcome will be reliable. A small variation in rating identifies agreement between raters while a large variation would indicate that the raters do not agree with each other. Intra-reliability occurs when one rater is responsible for all assessments, but the rater would need to be consistent when assessing on different occasions. In schools, departments moderate work to ensure the reliability of scores of classes taught by different teachers. In the current study, there was one rater who marked all the students' assessments. A marking scheme was devised and adhered to. Even then, it was difficult to ensure accurate marking as marking multiple assessments from 346 students could result in researcher exhaustion, and careless mistakes creeping in. The assessment were triple marked to ensure scores were consistent and therefore reliable.

Split-half reliability is a way to determine whether the instrument of measurement is correct, e.g., questions on a test: The participant should achieve similar scores if repeating the test. This is an internal measurement of consistency (homogeneity). "A high coefficient tells you that you can generalize from the score a person receives on one occasion to a score that person would receive if the test had been given at a different time" (Ary et al., 2018, p. 106). However, Ary et al. warn that the time between tests could affect reliability and this method should be used cautiously when measuring cognitive related tasks. Cronbach's α, determines the internal consistency of questions, i.e., whether there is consistency between items in a test. Reliability is demonstrated with a score of .7 or higher (Heale & Twycross, 2015). To determine the reliability of the current study, and to identify the precision and accuracy of the assessment questions which were based on the three curriculum areas of science, geography, and English, Cronbach's alpha was used. Results indicated–for combined immediate and delayed recall–that assessment questions were reliable (science α = .69, geography α = .76, English α = .65). The motivation questionnaire, adapted from Pintrich's (2003) Motivated Strategies for Learning Questionnaire, also proved to be highly reliable (α = .78), demonstrating a good internal consistency between questions (Field, 2014; Rovai, Baker, & Ponton, 2014).

A parallel form of reliability demonstrates the stability of an assessment. Test-retest is an example where a student should achieve a similar score on the same test given at different times in similar situations. Parallel-form reliability is similar, but instead of completing the same test, a slightly different test is provided. The content that the test is based on has not changed, just the wording on the assessment. Reliability is identified by a strong correlation between scores for each testing period (Heale & Twycross, 2015). The current study implemented this approach (see section 3.4.2.1.2.), and the correlation for mode of presentation and cognitive load for each curriculum subject was significant at the $p = .01$ level (Pearson's correlation r values were between 0.36 – 0.56).

Field (2014) acknowledged that sample size may have an impact on reliability, and determined that an acceptable sample size would be 10-15 participants per variable. In certain conditions, e.g., cognitive load and WM conditions in the heart task, the number of participants in the sample groups for WM, which ranged from 6-17, were not consistent (see Table 5.3., Table, 5.4., Table 5.5.), and would likely have reduced the reliability of the findings. Therefore, findings from the WM analyses must be viewed with caution.

3.7.2 Validity

The term external validity was first coined by Campbell (1957) who went on to equate the term external validity (which includes ecological validity) to generalisability. The findings of a research study should be generalisable to the population, and the testing sample group should be representative of the total population (Campbell & Stanley, 2015). Schmuckler (2001) identified that studies are often criticised for their lack of ecological validity, so much so, that the usefulness of the term comes into question. Ecological validity is not always possible. Stroebe, Gadenne and Nijstad (2018) criticised the belief that external validity identifies how generaliseable research findings can be and how representative the sample is of the general population–they acknowledged that this belief is erroneous. Stroebe et al. concurred with Mook (1983) who argued that " misplaced preoccupation with external validity can lead us to dismiss good research for which generalization to real life is not intended or meaningful" (p. 379). The aim of a piece of research may be to just explore a theory, or test a prediction, and therefore generalisability is not always essential. What is relevant is gaining valid information that supports or refutes a theory. The current study examined learning outcomes in relation to a theory, i.e., Meyer's (2005) principles of multimedia and redundancy, so in this respect generalisability was not a concern. However, in another respect the research was conducted in a realistic setting, i.e., a classroom in normal lesson times, and would be considered to have high ecological validity. Yet, students represented a non-randomised sample, which would be considered as lacking external validity. Schmuckler acknowledged "although care must be taken when generalizing across persons, settings, and responses, this is common with any experimental protocol, and simply reiterates the well-known importance of converging operations in research." (p. 420). It is hoped that other schools in a similar setting could gain from the findings of the current study.

As the current study was conducted over more than one testing time, experimental mortality, i.e., attrition, was an issue to consider. Internal validity could be reduced with the loss of participants between testing times (Ary et al., 2018). This factor was taken into account and only the particpants' data who were present at both times of testing was used for analyses.

Ary et al. (2018) identified the importance of content-related validity, in which the test questions reflect the content of the learning material in equal proportions. This can be ascertained through logical deduction and checking that the questions fully represent all the information contained the material. In the current study, questions fully represented the information contained in the data (e.g., see Appendices D and E).

Construct-related validity determines whether the test is a measure of the concept contained within a key theory, i.e., whether the assessment questions reflect Mayer's principles of multimedia and redundancy. Construct-related validity is intangible and difficult to measure, however, it can be determined via logical analysis. The known-groups technique allows the researcher to compare groups that differ within the concept being measured. An example in the current study would be the WM and motivation sample groups, where there were high WM and high motivation groups and low WM and low motivation groups. The assumption would be that the high level sample groups should achieve better than the low level sample groups. If a difference is found, one can conclude that the test is measuring the concepts it set out to measure–this can be determined via factor analysis. However, different variables should not correlate with each other even if they are consistent in their method of analysis (discriminant validity). A good premise would be that if the theory and data are well-matched, keep both the theory and data, otherwise determine which would need to be changed, either the theory or the assessment measurement (Ary, 2018).

3.8 Ethical Consideration

The importance of ethical consideration can never be overstated–it is crucial to the success of research as well as the reputation of empirical inquiry (Du, 2012). It has been succinctly put by Walliman and Buckler (2008) who noted that, "ethics is about moral principles and rules of conduct" (p. 30). Morally, researchers have an obligation to make research valuable to participants and help to improve a society's knowledge base through careful dissemination (Du, 2012; Hammersley & Traianou, 2012; Tobin, 2009). There is an obligation to protect the participants by continually evaluating the research (Strike, 2012).

Hammersley and Traianou (2012) recommended that participants should have a right to choose whether to participate in research. As such, and in order to gain informed consent, parents and participants were advised as to the nature of the research through a letter home and an assembly held at school for the participants. Parents signed a letter if they wished to opt out, and participants gave their consent in an initial questionnaire (see Appendix A). The school was also a participator in the current study, and consent was gained from the principal. The school was kept informed of the proceedings throughout the whole process.

Ethics to consider within the current study included researching minors (Cohen et al., 2017). With regards to psychological harm (Busher & James, 2012), the research was conducted in a familiar environment that did not differ from what the participants were used to in normal lessons. Participants were informed that their data would remain confidential, and were assured the results would not affect their school grades. Participants were given the opportunity to withdraw their data at any stage, and visit the school counsellor if they felt the need.

Confidentiality and anonymity are two important aspect of research (Cohen et al., 2017; Walliman & Buckler, 2008). Participants' identity and data were protected by assigning each individual a unique identifying number. In addition, all sensitive data were password protected. Any paper-based work was scanned electronically, after which the paper was shredded. The confidentiality and anonymity of the participants' data was maintained throughout the whole research process.

Ethnographic insider-research (when researchers are part of the group they are studying), and deception were considered. As the aim of the research was to identify optimal methods for learning, participants had to work on either a traditional medium (paper) or a more interactive medium (iPad®), within different cognitive load conditions. If the aim of the research was made known to the participants in advance, it may have influenced their approach to a task. This, in turn, could have had an impact on the outcome, and therefore the validity, of the research. Also, the researcher was familiar to the participants, which could be viewed as advantageous as the researcher had the ability to blend into the social setting, and knowing the participants could have engendered a greater level of honesty (Mercer, 2007). However, participants bring complex dispositions to the field of educational research–participants are never neutral and may have had different perceptions of the researcher (Rosenthal & Rosnow, 2009). Some participants may have wanted to please the researcher and their behaviour may have differed from the norm. These demand characteristics and ethnographical insider-research issues were counterbalanced by a single-blind methodology in which the participants did not have full access to the true nature of the research, i.e., they were not informed that cognitive load was being measured, or that there would be a second delayed recall assessment (for which they would have revised). Participants were informed of the purpose and general findings after the research took place. This could be construed as deception, but in this case it was necessary for the validity of the research, and supported by Rosenthal and Rosnow who stated that covert strategies, as well as other types of deception, are almost certainly unavoidable.

Sample material and ethical considerations were submitted to the university board of ethics, and approval was gained from the board in advance of the research taking place (see Appendix A).

3.9 Chapter Summary

Critical realism and activity theory paradigms were adopted in the current study (Kaptelinin & Nardi, 2006; Nardi, 1996; Sayer, 1992; Speed, 1991). Quantitative data analysis identified including the means and standard deviations to determine normal distribution, as well as ANOVAs to determine the main effects and interactions and between multiple variables (Field, 2014).

Students, as participants, prompted careful ethical consideration. Measures included procedures to prevent psychological harm and maintain confidentiality (Cohen et al., 2017; Walliman & Buckler, 2008). Ethical approval was gained from the university and a pilot study was conducted to identify if any improvements were necessary to the material and process. Resource design was identified as a factor that could affect learning, and resources were selected, or created, with particular design criteria in mind (Abeysekera & Dawson, 2015; Falloon, 2013).

Chapter 4 and Chapter 5 contain the statistical analyses of data collected over two testing times (immediate recall, delayed recall) in each of three tasks (heart, map, poem), and a discussion of the findings. Chapter 6 determine answers to the research questions, identifies limitations of the current study, discussions implications of the findings and identifies future research.

The research questions were:

- Does mode of presentation (paper, M-technology) affect learning outcome?
- Does cognitive load affect learning outcome?
- Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?

Chapter 4. Mode of Presentation Results

This chapter contains an analysis of the effect of mode of presentation (paper vs. M-technology) on participants' learning in separate tasks relating to three subject disciplines: science–topic heart; geography–topic map; and English–topic poem. Participants were assessed within an hour at the end of the learning phase, and again after a delay of four to six weeks, providing both immediate and delayed scores. Data were analysed to determine an answer to the research questions, "Does mode of presentation (paper, M-technology) affect learning outcome?" and "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?" The chapter begins with a description of the characteristics of participants (age, gender, cognitive ability, home language) then provides an initial exploration of data involving descriptive statistics, i.e., means and standard deviations. The between participation IVs were mode of presentation (paper, M-technology) and individual differences (gender, WM, motivation), the within participant IVs were testing time (immediate testing time, delayed testing time), and the DV was the recall scores. The covariate was prior knowledge. Inferential statistical analyses (ANOVA, ANCOVA) are reported to determine the extent to which the various IVs accounted for unique variance in recall. A summary of findings for each curriculum topic (heart, map, poem) can be found in Table 4.18.

4.1 General Characteristics of Participants

4.1.1 Participants' Age and Gender

Data were imported into SPSS® and initial analysis, which provided descriptive statistics of the age and gender of participants for each curriculum topic in each mode of presentation (paper, M-technology), can be viewed in Table 4.1. Age was largely consistent between sample groups and across curriculum topics. Sample sizes were not always equal due to participant absences, technology issues, and multiple independent variables, e.g., mode of presentation (paper, M-technology) and WM (high, low), being researched.

Table 4.1.

Mean Age and Standard Deviation of Participants for Mode of Presentation (Paper, M-technology) Analysis, with Data for Each Curriculum Topic (Heart, Map, Poem) and Gender (Males, Females) Presented Separately

		Paper			M-technology		
Task	*N*	Age *M*	Males N	Females N	Age *M*	Males N	Females N
Heart	211	12.99 (0.91)	54	54	13.10 (0.99)	52	51
Map	196	12.89 (0.89)	45	46	13.18 (0.98)	57	48
Poem	212	13.07 (0.98)	51	47	13.05 (0.92)	50	64

Note. Standard deviations are in parentheses.

4.1.2 Participants' Cognitive Ability

Cognitive Abilities Test (CAT) mean scores were calculated for each sample group in each curriculum topic to determine the mental capability of participants. This was to ensure that any differences found between the modes of presentation were not derived by the fact that there were differences in levels of ability between groups. Levels of cognitive ability were similar across the different sample groups (see Table 4.2.).

Table 4.2.

Means and Standard Deviations of CAT Scores (Cognitive Abilities Test) for Participants in Each Curriculum Topic (Heart, Map, Poem) with Mode of Presentation (Paper, M-technology) Presented Separately

	Paper		M-Technology	
Task	*N*	CAT *M*	*N*	CAT *M*
Heart	108	112.41 (8.91)	103	113.18 (9.04)
Map	91	114.27 (9.47)	105	113.62 (9.01)
Poem	98	113.61 (9.64)	114	113.14 (8.88)

Note. CAT *M* = Cognitive Abilities Test mean scores. Standard deviations are in parentheses.

4.1.3 Participants' Home Language

The research school was multi-national with over 50 ethnic groups represented. As the research materials were written/spoken in English, it was important to determine the home language of participants taking part. Within the total sample population (N = 346), 67% of participants' home language was English, which meant that for 33% of participants, English was a second language.

4.2 Descriptive Statistics for Each Sample Group in Each Curriculum Topic

In this section initial data analyses determined the means and standard deviations of immediate and delayed recall scores for mode of presentation (paper, M-technology) in each curriculum topic (heart, map, poem) for each sample group. CAT means and standard deviations are also provided. The individual difference variables were gender (male, female), WM scores (high, low), and motivation scores (high, low). High and low scores were determined by calculating the median–scores above the median were considered high scores, scores equal to or below the median were considered low scores. If groups varied, this could help to explain differences in outcome when answering the research questions, “Does mode of presentation (paper, M-technology) affect learning outcome?” and “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?” The results can be viewed in Table 4.3. (heart topic), Table 4.4. (map topic), and Table 4.5. (poem topic).

4.2.1 Heart Topic Descriptive Statistics

The means and standard deviations of recall scores (immediate, delayed) and CAT (Cognitive Abilities Test) scores for mode of presentation (paper, M-technology) were calculated for the heart curriculum topic, with scores for gender, WM, and motivation presented separately (see Table 4.3.).

Table 4.3.

Means and Standard Deviations of Recall Scores (Immediate, Delayed) and CAT (Cognitive Abilities Test) Scores for Mode of Presentation (Paper, M-technology) in the Heart Curriculum Topic with Scores for Gender, WM, and Motivation Presented Separately

			CAT Score		Immediate Recall Score		Delayed Recall Score	
	Variables	N	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
	Paper	108	112.41	(8.91)	10.49	(2.64)	5.72	(2.36)
	M-Technology	103	113.18	(9.04)	10.31	(2.58)	5.18	(2.86)
Gender	P Males	54	113.39	(7.79)	10.70	(2.50)	5.85	(2.35)
	P Females	54	111.43	(9.89)	10.28	(2.77)	5.59	(2.39)
	M-tech Males	52	112.98	(8.51)	9.77	(2.42)	5.04	(3.01)
	M-tech Females	51	113.39	(9.63)	10.86	(2.64)	5.33	(2.86)
WM	P WM Low Score	17	106.94	(9.87)	8.94	(3.17)	5.12	(2.67)
	P WM High Score	22	114.00	(7.56)	10.59	(2.70)	6.05	(2.54)
	M-tech WM Low Score	31	111.03	(8.56)	9.29	(2.24)	4.77	(2.94)
	M-tech WM High Score	25	115.24	(10.31)	11.04	(2.88)	5.64	(2.77)
Motivation	P Mot Low Score	56	111.88	(8.56)	10.41	(2.70)	5.46	(2.09)
	P Mot High Score	52	112.98	(9.32)	10.58	(2.59)	6.00	(2.62)
	M-tech Mot Low Score	61	113.10	(8.45)	9.84	(2.71)	4.36	(2.58)
	M-tech Mot High Score	42	113.31	(9.93)	11.00	(2.23)	6.38	(2.86)

Note. P = paper, Tech = M-technology, WM = working memory, and Mot = motivation. CAT = Cognitive Abilities Test scores.

4.2.2 Map Topic Means Descriptive Statistics

The means and standard deviations of recall scores (immediate, delayed) and CAT (Cognitive Abilities Test) scores for mode of presentation (paper, M-technology) in the map curriculum topic, with scores for gender, WM, and motivation presented separately (see Table 4.4.).

Table 4.4.

Means and Standard Deviations of Recall Scores (Immediate, Delayed) and CAT (Cognitive Abilities Test) Scores for Mode of Presentation (Paper, M-technology) in the Map Curriculum Topic with Scores for Gender, WM, and Motivation Presented Separately

Individual Differences	Variables	N	CAT Score *M*	CAT Score *(SD*	Immediate Recall Score *M*	Immediate Recall Score *(SD)*	Delayed Recall Score *M*	Delayed Recall Score *(SD)*
	Paper	91	114.27	(9.47)	4.63	(2.57)	1.96	(1.31)
	M-tech	103	113.62	(9.01)	4.66	(2.66)	2.00	(1.41)
Gender	P Males	45	115.60	(7.93)	3.91	(2.29)	1.84	(1.31)
	P Females	46	112.98	(10.70)	5.33	(2.66)	2.07	(1.31)
	M-tech Males	55	114.16	(8.89)	4.47	(2.82)	2.02	(1.33)
	M-tech Females	48	113.00	(9.21)	4.88	(2.46)	2.06	(1.73)
WM	P WM Low Score	23	112.48	(9.93)	4.04	(2.08)	1.70	(1.26)
	P WM High Score	27	115.63	(10.55)	4.44	(2.08)	2.15	(1.35)
	M-tech WM Low Score	27	110.44	(8.72)	4.00	(2.25)	1.81	(1.55)
	M-tech WM High Score	22	118.59	(8.18)	5.27	(3.33)	2.14	(1.52)
Motivation	P Mot Low Score	48	114.33	(9.90)	4.27	(2.57)	1.83	(1.31)
	P Mot High Score	43	114.21	(9.09)	5.02	(2.54)	2.09	(1.31)
	M-tech Mot Low Score	64	113.58	(8.56)	4.48	(2.64)	2.02	(1.24)
	M-tech Mot High Score	39	113.69	(9.82)	4.95	(2.70)	1.97	(1.68)

Note. P = paper, Tech = M-technology, WM = working memory, and Mot = motivation. CAT = Cognitive Abilities Test scores.

4.2.3 Poem Topic Means Descriptive Statistics

The means and standard deviations of recall scores (immediate, delayed) and CAT scores (Cognitive Abilities Test) for mode of presentation (paper, M-technology) in the poem curriculum topic, with scores for gender, WM, and motivation presented separately (see Table 4.5.).

Table 4.5.

Means and Standard Deviations of Recall Scores (Immediate, Delayed) and CAT Scores (Cognitive Abilities Test) for Mode of Presentation (Paper, M-technology) in the Poem Curriculum Topic with Scores for Gender, WM, and Motivation Presented Separately

			CAT Score		Immediate Recall Score		Delayed Recall Score	
	Variables	N	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
	Paper	98	113.61	(9.64)	9.61	(1.77)	8.72	(2.17)
	M-Tech	114	113.14	(8.88)	9.46	(1.58)	8.97	(1.96)
Gender	P Males	51	114.33	(9.08)	9.51	(1.58)	8.59	(1.91)
	P Females	47	112.83	(10.24)	9.72	(1.98)	8.91	(2.33)
	M-tech Males	50	114.08	(8.04)	9.44	(1.57)	8.76	(2.14)
	M-tech Females	64	112.41	(9.49)	9.47	(1.60)	9.13	(1.81)
WM	P WM Low Score	21	114.48	(8.42)	9.14	(1.80)	9.24	(1.84)
	P WM High Score	21	115.05	(10.83)	10.05	(1.56)	8.76	(2.26)
	M-tech WM Low Score	22	106.82	(8.57)	8.55	(1.30)	8.23	(2.58)
	M-tech WM High Score	25	116.28	(8.35)	9.76	(1.69)	9.45	(1.73)
Motivation	P Mot Low Score	53	113.17	(8.93)	9.62	(1.63)	8.83	(2.01)
	P Mot High Score	45	114.13	(10.48)	9.60	(1.95)	8.60	(2.36)
	M-tech Mot Low Score	66	114.61	(8.67)	9.29	(1.57)	8.79	(1.84)
	M-tech Mot High Score	48	111.13	(8.87)	9.69	(1.59)	9.21	(2.11)

Note. P = paper, Tech = M-technology, WM = working memory, and Mot = motivation. CAT = Cognitive Abilities Test scores.

4.3 Group Comparisons

Inferential statistical analyses were carried out to investigate the effect of mode of presentation (paper, M-technology), individual differences (gender, motivation, WM), and recall (immediate, delayed) were undertaken to determine an answer to the research questions "Does mode of presentation (paper, M-technology) affect learning outcome?" and "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?" Inferential statistical analyses (ANOVA, ANCOVA) were reported to test the differences between group means. A summary of the levels of significance for each curriculum topic can be viewed in Table 4.18.

4.3.1 Does Mode of Presentation have an Impact on Learning?

Main effects and interactions were explored for mode of presentation (paper, M-technology) across three separate curriculum topics (heart, map, poem) in two testing times (immediate recall, delayed recall). Statistical analyses were conducted to assess the impact of mode of presentation on time of recall–separate analysis was conducted for each curriculum topic. To determine individuals existing knowledge, a prior knowledge assessment was completed by participants in the heart and poem curriculum topics. Analysis of covariance (ANCOVA) was performed on curriculum topics in which prior knowledge was established (heart, poem), to determine whether within-group error variances existed, and analysis of variance (ANOVA) when no prior knowledge existed (map). The means and standard deviations of recall scores are presented in separate tables for each curriculum topic (Table 4.3., heart, Table 4.4., map, Table 4.5., poem). In instances where data violated the assumptions on which parametric tests are based, the non-parametric Mann-Whitney test was conducted. When parametric and non-parametric results concurred, parametric results were reported; when parametric and non-parametric results did not concur, the non-parametric results were reported. The between participant variables were mode of presentation (paper, M-technology) and individual differences (gender, WM, motivation), the within participant variable was time of recall (immediate, delayed), and the covariate for the ANCOVA analysis was prior knowledge.

4.3.1.1 Heart curriculum topic

The heart curriculum topic data were analysed using analysis of covariance (ANCOVA). The number of participants (male, female), mean recall scores, mean age, and mean CAT scores (with standard deviations), can be viewed in Table 4.1., Table 4.2., and Table 4.3., respectively. In certain instances data violated the assumptions of normality. In delayed recall Levene's test of equality of error variances was significant, p = .021, so both parametric and non-parametric Mann-Whitney tests were performed on the data. Results concurred so parametric results were reported.

4.3.1.1.1 Analysis of covariance for the heart curriculum topic

Results of the factori al ANCOVA analysis are summarised in Table 4.6. The means and standard deviations are presented separately in Table 4.3.

Table 4.6.

ANCOVA Main Effects and Interactions in the Heart Topic for Mode of Presentation (Paper, M-technology) in Two Testing Times (Immediate, Delayed) with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	227.31	1, 208	4.22	53.88 ***
Time * Prior Knowledge	13.73	1, 208	4.22	3.25
Time * Mode of Presentation	3.86	1, 208	4.22	.92
Prior Knowledge	259.47	1, 208	8.19	31.68 ***
Mode of Presentation	9.74	1, 208	8.19	1.19

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.30. **p* < .05. ***p* < .01. ****p* < .001.

ANCOVA analysis determined there were significant main effects of time of testing, $F_{(1, 208)} = 53.88$, $p < .001$, confirming students retained more information in the short term (immediate testing time), $M = 10.40$, $SE = .17$, than in the longer term (delayed testing time), $M = 5.45$, $SE = .18$. The covariate prior knowledge was significant, $F_{(1, 208)} = 31.68$, $p < .001$, which indicated prior knowledge had an influence on learning. To gain a fuller understanding of the nature of the covariate effect, additional correlation analysis was carried out, which determined that the direction of the correlation was positive, indicating that as prior knowledge scores scores increased, so did recall, $r = .38$, $p < .01$. The interaction between mode of presentation and time was not significant, so no further analysis was necessary.

4.3.1.1.2 Heart curriculum topic summary

In answer to the research question, “Does mode of presentation affect learning outcome?” analysis confirmed that the mode of presentation did not modulate the effects of recall in the science curriculum topic. No significant difference was found to students’ recall scores of the heart when data were presented on paper or M-technology media within an hour of being presented with the material, as well as four to six weeks later. Findings were in line with CAT scores, i.e., the cognitive ability, for each of the sample groups (see Table 4.3.).

4.3.1.2 Map curriculum topic

The map curriculum topic data were analysed using analysis of variance (ANOVA). The number of participants (male, female), mean recall scores, mean age, and mean CAT scores (with their respective standard deviations), can be viewed in Table 4.1.,

Table 4.2., and Table 4.4. In certain instances data violated the assumptions of normality, so both parametric and non-parametric tests were performed on the data. Non-parametric Mann-Whitney test results concurred with parametric analysis, so parametric results were reported.

4.3.1.2.1 Analysis of variance for the map curriculum topic

Results of the factorial ANOVA analysis are summarised in Table 4.7. Means and standard deviations are presented separately in Table 4.4.

Table 4.7.

ANOVA Main Effects and Interactions in the Map Topic for Mode of Presentation (Paper, M-technology) in Two Testing Times (Immediate, Delayed)

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	649.93	1, 192	2.43	267.55 ***
Time * Mode of Presentation	0.01	1, 192	2.43	0.01
Mode of Presentation	0.19	1, 192	6.91	0.03

*$p < .05$. **$p < .01$. ***$p < .001$.

ANOVA analysis determined there was a significant effect of time of testing, $F_{(1, 192)} = 267.55$, $p < .001$, confirming that students recalled more information about the map in the short term (immediate testing time), $M = 4.64$, $SE = .19$, than in the longer term (delayed testing time), $M = 2.05$, $SE = .11$. The interaction between mode of presentation and time was not significant, so no further analysis was necessary.

4.3.1.2.2 Map curriculum topic summary

In answer to the research question, "Does mode of presentation affect learning outcome?" analysis confirmed the mode of presentation did not modulate the effects of recall in the geography curriculum topic. Investigation revealed there was no significant difference in learning outcome between students who were presented with information on paper or M-technology media in a geography topic within an hour of being presented with the material, as well as four to six weeks later. Findings were in line with cognitive ability scores for each sample group (see Table 4.2.). Overall performance was low for the map curriculum topic in both testing times in both mode of presentation conditions (see Table 4.4).

4.3.1.3 Poem curriculum topic

The poem curriculum topic data were analysed using analysis of covariance (ANCOVA). The number of participants (male, female), mean recall scores (immediate, delayed), mean age, and mean CAT scores (with their standard deviations), can be viewed in Table 4.1., Table 4.2., and Table 4.5. respectively. Data violated the assumptions of normality, and extreme outliers, N = 8, were Winsorised. Both parametric and non-parametric tests were performed on the data. Non-parametric Mann-Whitney test results concurred with the parametric analysis, so parametric results were reported.

4.3.1.3.1 Analysis of covariance for the poem curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 4.8. Means and standard deviations are presented separately in Table 4.5.

Table 4.8.

ANCOVA Main Effects and Interactions in the Poem Topic for Mode of Presentation (Paper, M-technology) in Two Testing Times (Immediate, Delayed) with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	45.98	1, 209	1.86	24.84 ***
Time * Prior Knowledge	2.39	1, 209	1.86	1.29
Time * Mode of Presentation	3.46	1, 209	1.86	1.87
Prior Knowledge	11.54	1, 209	4.48	2.58
Mode of Presentation	0.00	1, 209	4.48	0.00

Note. Covariates appearing in the model are evaluated at the following values: Poem Prior Knowledge Raw Score /6 = .27

ANCOVA analysis determined there was a significant effect of time of testing, $F_{(1, 209)} = 24.84$, $p < .001$, confirming that students recalled more information about the poem in the short term (immediate testing time), $M = 9.55$, $SE = .11$, than in the longer term (delayed testing time), $M = 8.90$, $SE = .13$. The interaction between mode of presentation and time was not significant, so no further analysis was necessary.

4.3.1.3.2 Poem curriculum topic summary

In answer to the research question, “Does mode of presentation affect learning outcome?” analysis confirmed the mode of presentation did not modulate the effects of recall in the poem curriculum topic. Investigation revealed there was no significant difference in recall scores between students who were presented with information on paper or M-technology media, within an hour of being presented with the material as well as four to six weeks later. Findings were in line with the cognitive ability scores for each sample group (see Table 4.5.).

4.3.2 Does Gender Moderate the Effect of Mode of Presentation?

Main effects and interactions were explored for gender (males, females) and mode of presentation conditions (paper, M-technology) across three separate curriculum topics (heart, map, poem) in two testing times (immediate, delayed), to determine an answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?" The between participant variables were mode of presentation and gender, the within participant variable was time of recall, and the covariate for the ANCOVA analysis was prior knowledge.

4.3.2.1 Heart curriculum topic

The heart curriculum topic data were analysed using analysis of covariance (ANCOVA). The number of participants, means and standard deviations of recall scores, as well as CAT mean scores for each sample group, can be viewed in Table 4.3. In certain instances data violated the assumptions of normality, so both parametric and Mann-Whitney non-parametric tests were performed on the data. Results concurred, so parametric results were reported.

4.3.2.1.1 Analysis of covariance for the heart curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 4.9. Means and standard deviations are presented in Table 4.3.

Table 4.9.

ANCOVA Main Effects and Interactions in the Heart Topic for Mode of Presentation (Paper, M-technology) and Gender (Male, Female) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	222.78	1, 206	4.21	52.88 ***
Time * Prior Knowledge	14.86	1, 206	4.21	3.53
Time * Mode of Presentation	3.97	1, 206	4.21	0.94
Time * Gender	3.16	1, 206	4.21	0.75
Prior Knowledge	268.04	1, 206	8.08	33.19 ***
Mode of Presentation	9.43	1, 206	8.08	1.17
Gender	6.06	1, 206	8.08	0.75
Mode of Presentation * Gender	34.36	1, 206	8.08	4.25 *
Time * Mode of Presentation * Gender	6.78	1, 206	4.21	1.61

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.30. *p < .05. **p < .01. ***p < .001.

ANCOVA analysis determined there was a significant effect of time of testing, $F_{(1, 206)} = 52.88$, $p < .001$, confirming that students recalled more information in the short term (immediate testing time), $M = 10.41$, $SE = .17$, than in the longer term (delayed testing time), $M = 5.46$, $SE = .18$. The covariate prior knowledge was significant, $F_{(1, 206)} = 33.19$, $p < .001$, which indicated prior knowledge had an influence on learning. To gain a fuller understanding of the nature of the covariate effect, additional correlation analysis was carried out, which determined that the direction of the correlation was positive, indicating that as prior knowledge scores scores increased, so did recall, $r = .38$, $p < .01$. The interaction between gender and mode of presentation was significant, $F_{(1, 206)} = 4.25$, $p < .05$, which confirmed the effect of gender was not consistent across the different modes of presentation and the effect of mode of presentation was not consistent across genders.

Given the significant interaction between mode of presentation and gender, further simple effect tests, i.e., independent sample *t*-tests, were conducted. As the interaction between time and mode of presentation was not a significant factor, the data were collapsed across the two time points, and the mean of immediate recall and delayed recall scores was used for further analysis. The simple effect of gender was not significant for paper, $t_{(106)} = 0.84$, $p = .402$, or for M-technology, $t_{(101)} = -1.60$, $p = .113$. The simple effect of mode of presentation for girls was not significant, $t_{(103)} = -0.38$, $p = .709$, but the simple effect of mode of presentation for boys was significant, $t_{(104)} = 2.15$, $p = .034$, demonstrating that in the heart topic boys scored higher in the paper condition, $M = 8.28$, $SD = 2.01$, than in the M-technology condition, $M = 7.40$, $SD = 2.18$. Investigation revealed that boys performed better using paper-based resources in a science topic than M-technology-based resources.

4.3.2.1.2 Heart curriculum topic summary

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?” investigation revealed gender and mode of presentation modulated the effect of recall in the heart curriculum topic. Analysis confirmed boys retained more information when learning about the heart from paper-based resources than from M-technology-based resources. An explanation could be there were a greater number of boys with a higher cognitive ability in the paper condition compared to the M-technology condition. However, according to cognitive ability mean scores (see Table 4.3.), similar levels of cognitive ability occurred within male paper-based and M-technology-based sample groups. Therefore, evidence has determined, in this instance, boys performed better in a science topic when learning from paper-based resources than M-technology-based resources.

4.3.2.2 Map curriculum topic

The map curriculum topic data were analysed using analysis of variance (ANOVA). The number of participants, means scores and standard deviations of recall scores, as well as CAT mean scores for each sample group, can be viewed in Table 4.4. In certain instances data violated the assumptions of normality, so both parametric and non-parametric tests were performed on the data. Non-parametric Mann-Whitney test results concurred with the parametric analysis, so parametric results were reported.

4.3.2.2.1 Analysis of variance for the map curriculum topic

Results of the factorial ANOVA analysis are summarised in Table 4.10. Means and standard deviations are presented separately in Table 4.4.

Table 4.10.

ANOVA Main Effects and Interactions in the Map Topic for Mode of Presentation (Paper, M-technology) and Gender (Male, Female) in Two Testing Times (Immediate, Delayed)

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	651.67	1, 190	2.37	275.34 ***
Time * Mode of Presentation	0.00	1, 190	2.37	0.00
Time * Gender	15.65	1, 190	2.37	6.61 *
Gender	24.66	1, 190	6.80	3.63
Mode of Presentation	0.28	1, 190	6.80	0.04
Mode of Presentation * Gender	13.73	1, 190	6.80	2.02
Time * Mode of Presentation * Gender	1.61	1, 190	2.37	0.68

*$p < .05$. **$p < .01$. ***$p < .001$.

ANOVA analysis determined there was a significant effect of time of testing, $F_{(1, 190)} = 275.34$, $p < .001$, confirming that students recalled more information about the map in the short term (immediate testing time), $M = 4.65$, $SE = .19$, than in the longer term (delayed testing time), $M = 2.05$, $SE = .11$. The interaction between time and gender was significant, $F_{(1, 190)} = 6.61$, $p < .05$,confirming that girls retained significantly more information, $M = 5.10$, $SE = .27$, in the short term (immediate testing time) than boys, $M = 4.19$, $SE = .26$. The interaction between mode of presentation and time was not significant, so no further analysis was necessary.

4.3.2.2.2 Map curriculum topic summary

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?” analysis confirmed the mode of presentation did not modulate the effect of recall in a geography curriculum topic. Investigation revealed there was no significant difference in recall scores between boys and girls when presented with information on paper or M-technology media, within an hour of being presented with the material as well as four to six weeks later. However, girls retained more information than boys in the short term (immediate testing time). Differences in cognitive ability may have accounted for these differences, but CAT scores for boys, $M = 114.21$, $SD = 8.54$, were higher than for girls, $M = 112.99$, $SD = 9.77$, confirming that girls retained more information than boys in the short term (immediate testing time), regardless of their cognitive capacity. Performance in the map curriculum topic was relatively low in both testing times for both boys and girls (see Table 4.4.).

4.3.2.3 The English poem curriculum topic

The poem curriculum topic data were analysed using analysis of covariance (ANCOVA). The number of participants, mean scores and standard deviations, and CAT mean scores for each sample group, can be viewed in Table 4.5. In certain instances data violated the assumptions of normality, so both parametric and Mann-Whitney non-parametric tests were performed on the data. Non-parametric test results concurred with parametric analysis, so parametric results were reported.

4.3.2.3.1 Analysis of covariance for the poem curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 4.11. Means and standard deviations are presented separately in Table 4.5.

Table 4.11.

ANCOVA Main Effects and Interactions in the Poem Topic for Mode of Presentation (Paper, M-technology) and Gender (Male, Female) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	61.91	1, 207	2.46	25.17 ***
Time * Prior Knowledge	3.34	1, 207	2.46	1.36
Time * Mode of Presentation	5.55	1, 207	2.46	2.26
Time * Gender	0.98	1, 207	2.46	0.40
Prior Knowledge	14.23	1, 207	5.25	2.71
Mode of Presentation	0.67	1, 207	5.25	0.13
Gender	4.87	1, 207	5.25	0.93
Mode of Presentation * Gender	0.01	1, 207	5.25	0.97
Time * Mode of Presentation * Gender	0.77	1, 207	2.46	0.31

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.30. *$p < .05$. **$p < .01$. ***$p < .001$.

ANOVA analysis determined there was a significant effect of time of testing, $F_{(1, 207)} = 25.17$, $p < .001$, confirming that students recalled more information in the short term (immediate testing time), $M = 9.54$, $SE = .12$, than in the longer term (delayed testing time), $M = 8.79$, $SE = .15$. There were no further significant main effects or interactions. Given the lack of significant interactions involving gender, no further analysis was necessary.

4.3.2.3.2 Poem curriculum topic summary

In answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?" investigation revealed gender did not appear to modulate the effects of mode of presentation or testing time in the English curriculum topic, which was consistent with CAT mean scores across groups (see Table 4.5.). Investigation confirmed there was no significant difference in recall between boys and girls who were presented with a poem on paper or M-technology media within an hour of being presented with the material as well as four to six weeks later

4.3.3 Does Working Memory Moderate the Effect of Mode of Presentation?

Main effects and interactions were explored for WM and mode of presentation conditions across three separate curriculum topics in two testing times, to determine an answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?" To determine high and low WM, scores above the median were classified as high WM, and scores equal to or below than the median were classified as low WM. As participation for the WM assessment was voluntary, data collection resulted in small, uneven, sample sizes. The between participant variables were mode of presentation and WM, the within participant variable was time of recall, and the covariate for the ANCOVA analysis was prior knowledge.

4.3.3.1 Heart curriculum topic

The heart curriculum topic data were analysed using analysis of covariance (ANCOVA). The sample included 95 participants aged 11-14 years, (M = 12 years 8 months, SD = 0.79). The WM data were derived from the performance of 39 participants (males high WM = 16, males low WM = 6, females high WM = 6, females low WM = 11) whose input was via paper, and 56 participants (males high WM = 15, males low WM = 12, females high WM = 10, females low WM = 19) whose input was via M-technology. In certain instances data violated the assumptions of normality, so both parametric and non-parametric tests were performed on the data. Non-parametric test results concurred with parametric analysis, so parametric results were reported.

4.3.3.1.1 Analysis of covariance for the heart curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 4.12. Means and standard deviations are presented separately in Table 4.3.

Table 4.12.

ANCOVA Main Effects and Interactions in the Heart Topic for Mode of Presentation (Paper, M-technology) and WM (High, Low) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	61.16	1, 90	4.71	12.99 **
Time * Prior Knowledge	12.19	1, 90	4.71	2.59
Time * Mode of Presentation	5.68	1, 90	4.71	1.21
Time * WM	3.32	1, 90	4.71	0.71
Prior Knowledge	104.08	1, 90	9.06	11.49 **
Mode of Presentation	0.30	1, 90	9.06	0.86
WM	37.74	1, 90	9.06	4.16 *
Mode of Presentation * WM	2.37	1, 90	9.06	0.26
Time * Mode of Presentation * WM	0.08	1, 90	4.71	0.02

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.30. **p* < .05. ***p* < .01. ****p* < .001.

ANCOVA analysis determined there was a significant main effect of time of testing, $F_{(1, 90)} = 12.99$, $p < .01$, confirming that students recalled more information in the short term (immediate testing time), $M = 9.96$, $SE = .26$, than in the longer term (delayed testing time), $M = 5.39$, $SE = .29$. The covariate prior knowledge was significant, $F_{(1, 90)} = 11.49$, $p < .01$, which indicated prior knowledge had an influence on learning. To gain a fuller understanding of the nature of the covariate effect, additional correlation analysis was carried out, which determined that the direction of the correlation was positive, indicating that as prior knowledge scores scores increased, so did recall, $r = .38$, $p < .01$. There was a significant main effect of WM, $F_{(1, 90)} = 4.16$, $p < .05$, confirming that students with high WM, $M = 8.14$, $SE = .32$, retained more information about the heart than students with low WM, $M = 7.20$, $SE = .33$. There were no significant interactions. Given the lack of significant interactions involving WM, no further analysis was necessary.

4.3.3.1.2 Heart curriculum topic summary

In answer to the research question "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?" investigation revealed WM did not appear to modulate the effects of mode of presentation or testing time in the science curriculum topic. High WM students performed better than low WM students in the heart topic, which was consistent with the cognitive ability mean scores (see Table 4.3.). Investigation confirmed there was no significant difference in recall between students with high and low WM when presented with information on paper or M-technology media within an hour of being presented with the material as well as four to six weeks later.

4.3.3.2 Map curriculum topic

The map curriculum topic data were analysed using analysis of variance (ANOVA). The sample included 99 participants aged 11-14 years, (*M* = 12 years 6 months, *SD* = 0.83). The WM data were derived from the performance of 50 participants (males high WM = 17, males low WM = 11, females high WM = 10, females low WM = 12) whose input was via paper, and 49 participants (males high WM = 18, males low WM = 11, females high WM = 4, females low WM = 16) whose input was via M-technology. In certain instances data violated the assumptions of normality. Levene's test of equality of variances was significant in immediate recall, *p* = .015, so both parametric and non-parametric tests were performed on the data. Non-parametric analysis concurred with parametric analysis, so parametric results were reported.

4.3.3.2.1 Analysis of variance for the map curriculum topic

Results of the factorial ANOVA analysis are summarised in Table 4.13. Means and standard deviations are presented separately in Table 4.4.

Table 4.13.

ANCOVA Main Effects and Interactions in the Map Topic for Mode of Presentation (Paper, M-technology) and WM (High, Low) in Two Testing Times (Immediate, Delayed)

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	284.30	1, 95	1.93	147.14 ***
Time * Mode of Presentation	0.73	1, 95	1.93	0.38
Time * WM	1.24	1, 95	1.93	0.64
WM	22.56	1, 95	6.78	3.33
Mode of Presentation	3.59	1, 95	6.78	0.53
Mode of Presentation * WM	2.25	1, 95	6.78	0.33
Time * Mode of Presentation * WM	2.41	1, 95	1.93	1.25

**p* < .05. ***p* < .01. ****p* < .001.

ANOVA analysis determined there was a significant effect of time of testing, $F_{(1, 95)} = 147.14$, $p < .001$, confirming that students recalled more information about the map in the short term (immediate testing time), $M = 4.44$, $SE = .25$, than in the longer term (delayed testing time), $M = 2.03$, $SE = .17$. There were no further significant main effects or interactions. Given the lack of significant interactions involving WM, no further analysis was necessary.

4.3.3.2.2 Map curriculum topic summary

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?” investigation confirmed WM did not modulate the effects of mode of presentation or testing time in the geography curriculum topic. Even though CAT mean scores for students in the high WM paper- and M-technology-based sample groups were higher than for students in the low WM paper- and M-technology-based sample groups, analysis confirmed there was no significant difference in recall between high and low WM students when presented with information on paper or M-technology media. The map scores for all groups in all conditions were relatively low (see Table 4.4.).

4.3.3.3 Poem curriculum topic

The poem curriculum topic data were analysed using analysis of covariance (ANCOVA). The sample included 89 participants aged 11-14 years, (M = 12 years 7 months, $SD = 0.78$). The WM data were derived from the performance of 23 participants (males = 23, females = 19) whose input was via paper, and 47 participants (males = 23, females = 24) whose input was via M-technology. In certain instances data violated the assumptions of normality. Non-parametric test results concurred with parametric analysis, so parametric results were reported.

4.3.3.3.1 Analysis of covariance for the poem curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 4.14. Means and standard deviations are presented separately in Table 4.5.

Table 4.14.

ANCOVA Main Effects and Interactions in the Poem Topic for Mode of Presentation (Paper, M-technology) and WM (High, Low) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	14.89	1, 84	2.42	6.15 *
Time * Prior Knowledge	4.18	1, 84	2.42	1.73
Time * Mode of Presentation	1.28	1, 84	2.42	0.53
Time * WM	7.40	1, 84	2.42	3.06
Prior Knowledge	2.33	1, 84	5.44	0.43
Mode of Presentation	3.38	1, 84	5.44	0.62
WM	18.48	1, 84	5.44	3.40
Mode of Presentation * WM	13.32	1, 84	5.44	2.45
Time * Mode of Presentation * WM	6.73	1, 84	2.42	2.78

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.30. *p < .05. **p < .01. ***p < .001.

ANCOVA analysis determined there was a significant effect of time of testing, $F_{(1, 84)} = 6.15$, $p < .05$, confirming that students recalled more information about the poem in the short term (immediate testing time), $M = 9.37$, $SE = .17$, than in the longer term (delayed testing time), $M = 8.86$, $SE = .24$. There were no other significant main effects or interactions. Given the lack of significant interaction between mode of presentation and WM, no further analysis was necessary.

4.3.3.3.2 Poem curriculum topic summary

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?” investigation revealed WM did not appear to modulate the effects of mode of presentation in the English curriculum topic. Analysis confirmed there was no significant difference in learning outcome between high and low WM students who were presented with information on paper or M-technology media in a poem topic within an hour of being presented with the material, as well as four to six weeks later.

4.3.4 Does Motivation Moderate the Effect of Mode of Presentation?

Main effects and interactions were explored for motivation and mode of presentation conditions across three separate curriculum topics in two testing times, to determine an answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?” To determine high and low motivation, scores above the median were classified as high motivation, and scores equal to or lower than the median were classified as low motivation. The between participant variables were mode of presentation and motivation, the within participant variable was time of recall, and the covariate for the ANCOVA analysis was prior knowledge.

4.3.4.1 Heart curriculum topic

The heart curriculum topic data were analysed using analysis of covariance (ANCOVA). The sample included 211 participants aged 11-14 years, (*M* = 13 years 4 months, *SD* = 0.95). The motivation data were derived from the performance of 108 participants (males high motivation = 26, males low motivation = 28, females high motivation = 26, females low motiviation = 28) whose input was via paper, and 103 participants (males high motivation = 23, males low motivation = 29, females high motivation = 19, females low motivation = 32) whose input was via M-technology. In certain instances data violated the assumptions of normality, and both parametric and non-parametric tests were performed on the data. Non-parametric test results concurred with parametric analysis, so parametric results were reported. The number of participants, as well as CAT mean scores for each condition, can be viewed in Table 4.3.

4.3.4.1.1 Analysis of covariance for the heart curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 4.15. Means and standard deviations are presented separately in Table 4.3.

Table 4.15.

ANCOVA Main Effects and Interactions in the Heart Topic for Mode of Presentation (Paper, M-technology) and Motivation (High, Low) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	203.97	1, 206	4.19	48.67 ***
Time * Prior Knowledge	16.98	1, 206	4.19	4.05 *
Time * Mode of Presentation	2.34	1, 206	4.19	0.56
Time * Motivation	12.26	1, 206	4.19	2.93
Prior Knowledge	219.63	1, 206	7.82	28.09 ***
Mode of Presentation	3.78	1, 206	7.82	0.48
Motivation	70.67	1, 206	7.82	9.04 *
Mode of Presentation * Motivation	26.12	1, 206	7.82	3.34
Time * Mode of Presentation * Motivation	2.44	1, 206	4.19	0.58

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.30.
**p* < .05. ***p* < .01. ****p* < .001.

ANCOVA analysis determined there was a significant main effect of time of testing, $F_{(1, 206)} = 48.67$, $p < .001$, confirming that students recalled more information about the heart within an hour of being presented with the learning material (immediate testing time), $M = 10.44$, $SE = .17$, than after four to six weeks (delayed testing time), $M = 5.54$, $SE = .17$. There was a significant main effect of prior knowledge, $F_{(1, 206)} = 28.09$, $p < .001$, confirming that prior knowledge influenced learning. To gain a fuller understanding of the nature of the covariate effect, additional correlation analysis was carried out, which determined that the direction of the correlation was positive, indicating that as prior knowledge scores scores increased, so did recall, $r = .38$, $p < .01$. Findings determined there was a significant main effect of motivation, $F_{(1, 206)} = 9.04$, $p < .05$, confirming that students with high motivation, $M = 8.41$, $SE = .21$, retained more information about the heart than students with low motivation, $M = 7.58$, $SE = .18$. Given the lack of significant interaction involving motivation and mode of presentation, no further analysis was necessary.

4.3.4.1.2 Heart curriculum topic summary

In answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?" investigation revealed that while motivation modulated overall performance, it did not modulate the effect of mode of presentation in the science curriculum topic. Students with high motivation performed better than students with low motivation, even though all groups had similar cognitive abilities (see Table 4.3.). However, this was not relative to mode of presentation. Analysis confirmed there was no significant difference in learning outcome between high and low motivation students who were presented with information on paper or M-technology media in a heart topic within an hour of being presented with the material, as well as four to six weeks later.

4.3.4.2 Map curriculum topic

The map curriculum topic data were analysed using analysis of variance (ANOVA). The sample included 194 participants aged 11-14 years, (M = 12 years 8 months, $SD = 0.95$). The motivation data were derived from the performance of 91 participants (males high motivation = 22, males low motivation = 23, females high motivation = 21, females low motivation = 25) whose input was via paper, and 103 participants (males high motivation = 20, males low motivation = 35, females high motivation = 19, females low motivation = 29) whose input was via M-technology. In certain instances data violated the assumptions of normality, so both parametric and non-parametric tests were performed on the data. Non-parametric test results concurred with parametric analysis, so parametric results were reported.

4.3.4.2.1 Analysis of variance for the map curriculum topic

Results of the factorial ANOVA analysis are summarised in Table 4.16. Means and standard deviations are presented separately in Table 4.4.

Table 4.16.

ANCOVA Main Effects and Interactions in the Map Topic for Mode of Presentation (Paper, M-technology) and Motivation (High, Low) in Two Testing Times (Immediate, Delayed)

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	645.33	1, 190	2.43	265.71 ***
Time * Mode of Presentation	0.00	1, 190	2.43	0.00
Time * Motivation	4.88	1, 190	2.43	2.01
Motivation	13.55	1, 190	6.91	1.96
Mode of Presentation	0.39	1, 190	6.91	0.06
Mode of Presentation * Motivation	1.08	1, 190	6.91	0.16
Time * Mode of Presentation * Motivation	0.13	1, 190	2.43	0.05

*$p < .05$. **$p < .01$. ***$p < .001$.

ANOVA analysis determined there was a significant effect of time of testing, $F_{(1, 190)} = 265.71$, $p < .001$, confirming that students recalled more information about the map in the short term (immediate testing time), $M = 4.68$, $SE = .19$, than in the longer term (delayed testing time), $M = 2.06$, $SE = .12$. There were no further significant main effects or interactions. Given the lack of significant interactions involving motivation, no further analysis was necessary.

4.3.4.2.2 Map curriculum topic summary

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?” investigation revealed that motivation did not modulate the effects of mode of presentation or testing time in the geography curriculum topic. Analysis confirmed there was no significant difference in learning outcome between high motivation students and low motivation students who were presented with information on paper or M-technology media in a map topic within an hour of being presented with the material, as well as four to six weeks later. The map scores for all groups in all conditions were relatively low (see Table 4.4.).

4.3.4.3 Poem curriculum topic

The poem curriculum topic data were analysed using analysis of covariance (ANCOVA). The sample included 212 participants aged 11-14 years, (*M* = 13 years 6 months, *SD* = 0.94). The motivation data were derived from the performance of 98 participants (males high motivation = 26, males low motivation = 25, females high motivation = 19, females low motivation = 28) whose input was via paper, and 114 participants (males high motivation = 22, males low motivation = 28, females high motivation = 26, females low motivation = 38) whose input was via M-technology. In certain instances data violated the assumptions of normality. Non-parametric test results concurred with parametric analysis, so parametric results were reported. The number of participants, as well as CAT mean scores for each condition, can be viewed in Table 4.5.

4.3.4.3.1 Analysis of covariance for the poem curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 4.17. Means and standard deviations are presented separately in Table 4.5.

Table 4.17.

ANCOVA Main Effects and Interactions in the Poem Topic for Mode of Presentation (Paper, M-technology) and Motivation (High, Low) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	61.13	1, 207	2.46	24.83 ***
Time * Prior Knowledge	3.57	1, 207	2.46	1.45
Time * Mode of Presentation	6.39	1, 207	2.46	2.60
Time * Motivation	0.75	1, 207	2.46	0.30
Prior Knowledge	15.29	1, 207	5.21	2.93
Mode of Presentation	1.97	1, 207	5.21	0.38
Motivation	1.05	1, 207	5.21	0.20
Mode of Presentation * Motivation	9.18	1, 207	5.21	1.76
Time * Mode of Presentation * Motivation	0.63	1, 207	2.46	0.26

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.30.
**p* < .05. ***p* < .01. ****p* < .001.

ANCOVA analysis determined there was a significant effect of time of testing, $F_{(1, 207)} = 24.83$, $p < .001$, confirming that students recalled more information about the poem in the short term (immediate testing time), $M = 9.55$, $SE = .12$, than in the longer term (delayed testing time), $M = 8.80$, $SE = .15$. There were no further significant main effects or interactions. Given the lack of significant interactions involving motivation, no further analysis was necessary.

4.3.4.3.2 Poem curriculum topic summary

In answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?" investigation revealed motivation did not appear to modulate the effects of mode of presentation or testing time in the English curriculum topic. Analysis confirmed there was no significant difference in learning outcome between high motivation students and low motivation students on paper or M-technology media in a poem topic within an hour of being presented with the material, as well as four to six weeks later.

4.3.5 Group Comparison Summary

Data were analysed separately for three curriculum topics (science heart, geography map, English poem) to determine answers to the research question, "Does mode of presentation affect learning outcome?" The additional factors of gender, WM, and motivation were explored to determine answers to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?" A summary of the levels of significance for each curriculum topic can be viewed in Table 4.18.

In answer to the research question, "Does mode of presentation affect learning outcome?" findings confirmed the mode of presentation did not modulate the effects of learning in each of the three curriculum topics. Results established there was no benefit to learning using M-technology-based resources, and no detriment to learning using paper-based resources.

To answer the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?" gender differences confirmed that boys performed better on paper-based resources than on M-technology-based resources in the heart curriculum topic. A significant effect of gender was found only in the heart curriculum topic, not in the map or poem curriculum topics. WM did not modulate the effect of mode of presentation at either time of testing in each of the three curriculum topics, i.e., there was no difference in performance between students with high WM and students with low WM using either paper- or M-technology-based resources–retention was consistent across both groups. There was a similar outcome for motivation.

Table 4.18.

Summary of Significant Main Effects and Interactions for Each Curriculum Topic (Heart, Map, Poem) for Mode of Presentation (Paper, M-technology) with Individual Differences (Gender, WM, motivation) Presented Separately

Variables	Simple Main Effects & Interactions	Heart Task	Map Task	Poem Task
	Time	***	***	***
	Prior Knowledge	***	-	-
	Mode of Presentation	-	-	-
	Time x Prior Knowledge	-	-	-
	Time x Mode of Presentation	-	-	-
Gender	Time	***	***	***
	Prior Knowledge	***	-	-
	Mode of Presentation	-	-	-
	Gender	-	-	-
	Time x Prior Knowledge	-	-	-
	Time x Mode of Presentation	-	-	-
	Time x Gender	-	*	-
	Mode of Presentation x Gender	*	-	-
	Time x Mode of Presentation x Gender	-	-	-
WM	Time	***	***	*
	Prior Knowledge	**	-	-
	Mode of Presentation	-	-	-
	WM	*	-	-
	Time x Prior Knowledge	-	-	-
	Time x Mode of Presentation	-	-	-
	Time x WM	-	-	-
	Mode of Presentation x WM	-	-	-
	Time x Mode of Presentation x WM	-	-	-
Motivation	Time	***	***	***
	Prior Knowledge	***	-	-
	Mode of Presentation	-	-	-
	Motivation	*	-	-
	Time x Prior Knowledge	*	-	-
	Time x Mode of Presentation	-	-	-
	Time x Motivation	-	-	-
	Mode of Presentation x Motivation	-	-	-
	Time x Mode of Presentation x Motivation	-	-	-

*$p < .05$. **$p < .01$. ***$p < .001$.

Performance was relatively low in the map curriculum topic, and students did not appear to retain the information they learnt, which indicated that the information was initially not strongly encoded, and/or the curriculum topic was too difficult. However, there is further dialog on this issue in the Discussion. In conclusion, findings confirmed no difference in learning outcome between paper- and M-technology-based resources in each of the three curriculum topics. Gender may need to be considered when identifying the best mode of presentation in science-based curriculum topics.

4.4 Specific Characteristics of Participants (Age–Years 7, 8, and 9)

Participants represented a wide age range of students (11-14 years) within KS3, and age could have had an impacted on recall. Analyses of variance were carried out to investigate whether any main effects or interactions existed between mode of presentation, year group (7, 8, 9), and recall in both testing times (immediate, delayed) in each of the three topic curriculum areas (heart, map, poem). Results can be viewed in Table 4.19.

Table 4.19.

Analyses of Variance Examining the Difference Between Mode of Presentation Scores (Paper, M-technology) in Two Testing Times (Immediate, Delayed) in Each of the Three Curriculum Areas (Heart, Map, Poem), with Data for Years 7, 8, and 9 Presented Separately.

Task	Year Grp	Mode of Presentation	*N*	Time 1 Immediate Recall *M*	*F*	*p*	Time 2 Delayed Recall *M*	*F*	*p*
Heart	7	Paper	39	9.08 (2.79)	1.22	.273	5.62 (2.56)	0.00	.982
		M-technology	32	9.31 (2.46)			5.25 (3.04)		
	8	Paper	41	11.44 (2.05)	2.37	.128	5.98 (2.23)	2.64	.108
		M-technology	36	10.69 (2.77)			5.08 (2.56)		
	9	Paper	28	11.07 (2.40)	0.40	.527	5.50 (2.32)	0.85	.362
		M-technology	35	10.83 (2.28)			5.23 (3.07)		
Map	7	Paper	32	4.50 (2.23)	0.33	.570	2.16 (1.32)	1.49	.227
		M-technology	23	4.13 (2.55)			1.74 (1.14)		
	8	Paper	43	4.60 (2.68)	0.40	.531	2.05 (1.76)	0.08	.782
		M-technology	35	5.00 (2.85)			1.94 (1.47)		
	9	Paper	16	4.94 (3.02)	0.12	.731	1.69 (1.25)	1.09	.301
		M-technology	45	4.67 (2.57)			2.09 (1.35)		
Poem	7	Paper	31	9.42 (1.69)	0.52	.472	9.32 (1.64)	0.14	.708
		M-technology	39	9.13 (1.64)			9.49 (2.03)		
	8	Paper	37	9.59 (1.98)	1.01	.318	8.41 (2.27)	0.10	.753
		M-technology	36	9.19 (1.43)			8.58 (2.27)		
	9	Paper	30	9.83 (1.62)	0.44	.512	8.50 (2.54)	0.85	.360
		M-technology	39	10.03 (1.53)			8.74 (1.68)		

Note. Year Grp = Year Group. Year 7–mean age 12.00, Year 8–mean age 13.09, Year 9–mean age 14.07. Heart score total out of 14, Map score out of 14, and Poem score out of 13. Standard deviations are in parentheses.

Analysis confirmed there was no significant difference in learning outcome between Year 7, 8, and 9 students who were presented with information on paper or M-technology media in each of the three curriculum topics (heart, map, poem) within an hour of being presented with the material, as well as four to six weeks later. Even though mean scores differed between year groups, there was no difference to learning outcome with regards to mode of presentation.[1]

4.5 Chapter Summary

Statistical analyses were conducted to determine anwers to the research questions, “Does mode of presentation (paper, M-technology) affect learning outcome?” and “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?” At the start of the chapter, the general characteristics of students were identified: The mean age and distribution of boys and girls was provided for each sample group (see Table 4.1., Table 4.2), and the means and standard deviations of student scores in each of the three curriculum topics (heart, map, poem) in immediate and delayed testing times (see Table 4.3., Table 4.4., Table 4.5.) were also presented. The distribution of students’ cognitive ability (CAT) scores across mode of presentation conditions was also provided. CAT scores identified the consistency of the cognitive capacity of students across experimental groups as inconsistent groups may have effected the findings. Analysis of variance was later performed on data from students in Year 7, 8, and 9 to determine whether recall was consistent across year groups, i.e., whether age effected learning outcomes (see Table 4.19.). There were no significant interactions between year groups and mode of presentation, which supported the selection of the sample group from KS3 taking part in the current study

In answer to the research question, “Does mode of presentation (paper, M-technology) affect learning outcome?” Group comparison analyses (ANCOVA/ANOVA) confirmed there was no significant difference to learning outcome in each of the three curriculum topics (heart, map, poem) between resources presented on paper or M-technology media. Findings showed that paper-based material is still a valid form of study and that M-technology is not detrimental to learning.

---

[1] As the research took place in an international school where 33% of participants’ home language was not English, it was important to determine whether language had an influence on learning outcome. Analyses confirmed one significant main effect involving language (see Appendix Q).

The individual differences of gender, WM, and motivation were considered as potential variables impacting on learning outcome. The relationship between these variables and mode of presentation learning outcome was explored to determine an answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation on learning outcome?” There was only one interaction between mode of presentation and the individual difference of gender that was significant. Analysis confirmed boys retained more information when learning about the heart from paper-based resources than from M-technology-based resources. Findings identified when preparing lessons, educators should consider the gender of students and plan accordingly.

Chapter 5. Cognitive Load Results

This chapter contains an analysis of the effect of cognitive load (text-only; text & graphics; graphics & audio; text, graphics & audio) on participants' learning. Participants completed separate tasks relating to three subject disciplines: science–topic heart; geography–topic map; and English–topic poem. Participants were assessed within an hour after the learning phase, and again after a delay of four to six weeks, providing both immediate and delayed scores. The data are the same set as reported in Chapter 4, but in this chapter analysis was conducted to determine an answer to the research question, "Does cognitive load affect learning outcome?" Analysis was also conducted separately for the factors of gender, WM, and motivation, to determine an answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of cognitive load on learning outcome?" The chapter begins with a description of the characteristics of participants (age, gender, cognitive ability, home language) then provides an initial exploration of data involving descriptive statistics, i.e., means and standard deviations. The between participation IVs were cognitive load (text-only; text & graphics; graphics & audio; text, graphics & audio) and individual differences (gender, WM, motivation), the within participant IVs were testing time (immediate testing time, delayed testing time), and the DV was the recall scores. The covariate was prior knowledge. Inferential statistical analyses (ANOVA, ANCOVA) are reported to determine the extent to which the various IVs accounted for unique variance in recall. A summary of findings for each curriculum topic (heart, map, poem) can be found in Table 5.20.

5.1 General Characteristics of Participants

5.1.1 Participants' Age and Gender

Data were imported into SPSS® and initial analysis, which provided descriptive statistics of the age and gender of participants for each curriculum topic in each cognitive load condition (text-only; text & graphics; graphics & audio; text, graphics & audio), can be viewed in Table 5.1. Sample sizes were not always equal due to participant absences, technology issues, and multiple independent variables being researched, e.g., four cognitive load conditions and WM (high, low).

Table 5.1.

Mean Age and Standard Deviation of Participants for Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) with Data for Each Curriculum Topic (Heart, Map, Poem) and Gender (Males, Females) Presented Separately

		Text-only			Text & Graphics			Graphics & Audio			Text, Graphics & Audio		
Task	*N*	Age *Mean*	M	F	Age *Mean*	M	F	Age *Mean*	M	F	Age *Mean*	M	F
Heart	203	13.33 (1.05)	36	32	12.65 (0.69)	16	19	13.03 (0.96)	25	24	13.15 (0.95)	23	28
Map	198	13.44 (0.95)	30	28	12.91 (0.95)	25	20	12.91 (1.01)	21	26	13.01 (0.90)	25	23
Poem	208	13.04 (0.94)	25	33	13.07 (0.90)	25	31	12.91 (0.90)	25	21	13.10 (1.08)	30	18

Note. M = number of male participants, F = number of female participants. Standard deviations are in parentheses.

5.1.2 Participants' Cognitive Ability

Cognitive Abilities Test (CAT) mean scores were calculated for sample groups in each curriculum topic to determine the mental capability of participants. This was to ensure that any differences found between the cognitive loads were not derived by the fact that there were differences in levels of ability between groups. Levels of cognitive ability were similar across the different sample groups (see Table 5.2.).

Table 5.2.

Means and Standard Deviations of CAT Scores (Cognitive Abilities Test) for Participants Completing Each Task (Heart, Map, Poem) for All Cognitive Load Conditions (Text-only, Text & Graphics, Graphics & Audio, Text, Graphics & Audio)

	Text-Only		Text & Graphics		Graphics & Audio		Text, Graphics & Audio	
Task	*N*	CAT *M*	*N*	CAT *M*	*N*	CAT *M*	*N*	CAT *M*
Heart	68	113.03 (8.51)	35	113.49 (10.11)	49	114.55 (9.56)	51	114.12 (9.05)
Map	58	114.67 (8.85)	45	112.27 (9.13)	47	113.79 (9.03)	48	113.73 (9.46)
Poem	58	112.50 (8.50)	56	113.80 (9.30)	46	112.26 (9.40)	48	112.42 (8.15)

Note. CAT *M* = Cognitive Abilities Test mean score. Standard deviations are in parentheses.

5.1.3 Participants' Home Language

The research school was multi-national with over 50 ethnic groups represented. As the research materials were written/spoken in English, it was important to determine the home language of participants taking part, as language could have had an impact on learning. Within the total sample population (N = 346), 67% of participants' home language was English, which meant that for 33% of participants, English was a second language.

5.2 Descriptive Statistics for Each Sample Group in Each Curriculum Topic

In this section initial data analyses determined the means and standard deviations of immediate and delayed recall scores for cognitive load (text-only; text & graphics; graphics & audio; text, graphics & audio) in each curriculum topic (heart, map, poem) for each sample group. CAT means and standard deviations are also provided. The individual difference variables were gender (male, female), WM scores (high, low), and motivation scores (high, low). High and low scores were determined by calculating the median–scores above the median were considered high scores, scores equal to or below the median were considered low scores. If groups varied, this could help to explain differences in outcome when answering the research questions, “Does cognitive load affect learning outcome?” and “Do individual differences (gender, WM, motivation) moderate the effect of cognitive load on learning outcome?” The results can be viewed in Table 5.3. (heart topic), Table 5.4. (map topic), and Table 5.5. (poem topic).

5.2.1 Heart Topic Descriptive Statistics

Descriptive statistics of recall scores (immediate, delayed) and CAT (Cognitive Abilities Test) scores for each cognitive load condition (text-only; text & graphics; graphics & audio; text, graphics & audio) the heart curriculum topic with scores for gender, WM, and motivation presented separately (see Table 5.3.).

Table 5.3.

Means and Standard Deviations of Recall Scores (Immediate, Delayed) and CAT Scores (Cognitive Abilities Test) for Cognitive Load (Text-Only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) in the Heart Curriculum Topic with Scores for Gender, WM, and Motivation presented separately

Individual Differences	Cognitive Load Variables	N	CAT Score *M*	CAT Score *(SD)*	Time 1 *M*	Time 1 *(SD)*	Time 2 *M*	Time 2 *(SD)*
	Text-Only	68	113.03	(8.51)	10.47	(2.55)	5.06	(2.94)
	Text & Graphics	35	113.49	(10.11)	10.00	(2.65)	5.43	(2.74)
	Graphics & Audio	49	114.55	(9.56)	10.29	(1.94)	5.61	(2.66)
	Text, Graphics & Audio	51	114.12	(9.05)	10.20	(2.51)	6.39	(3.03)
Gender								
Males	Text-Only	36	112.50	(8.79)	9.92	(2.62)	4.89	(3.07)
	Text & Graphics	16	114.06	(8.01)	9.44	(1.93)	5.38	(2.94)
	Graphics & Audio	25	114.44	(9.30)	10.28	(1.95)	5.28	(2.81)
	Text, Graphics & Audio	23	115.65	(9.43)	10.13	(2.55)	6.91	(2.86)
Female	Text-Only	32	113.63	(8.29)	11.09	(2.35)	5.25	(2.82)
	Text & Graphics	19	113.00	(11.79)	10.47	(3.10)	5.47	(2.64)
	Graphics & Audio	24	114.67	(10.02)	10.29	(1.97)	5.96	(2.51)
	Text, Graphics & Audio	28	112.86	(8.69)	10.25	(2.53)	5.96	(3.16)
WM								
Low	Text-Only	17	108.76	(7.82)	9.47	(2.48)	4.35	(2.89)
	Text & Graphics	14	113.79	(8.86)	9.07	(1.98)	5.29	(3.02)
	Graphics & Audio	8	115.63	(7.96)	10.13	(1.46)	4.88	(3.14)
	Text, Graphics & Audio	13	111.23	(9.08)	8.54	(2.40)	4.92	(2.99)
High	Text-Only	13	115.23	(8.69)	11.00	(2.65)	5.62	(3.07)
	Text & Graphics	12	115.25	(12.23)	11.08	(3.23)	5.67	(2.54)
	Graphics & Audio	6	118.33	(7.23)	9.33	(1.75)	4.67	(3.20)
	Text, Graphics & Audio	11	120.18	(8.18)	9.91	(2.02)	7.64	(2.06)
Motivation								
Low	Text-Only	43	113.33	(7.05)	10.00	(2.65)	4.14	(2.48)
	Text & Graphics	18	112.56	(11.35)	9.44	(2.90)	4.89	(2.78)
	Graphics & Audio	29	113.59	(7.46)	10.03	(1.99)	5.28	(2.39)
	Text, Graphics & Audio	31	115.71	(9.39)	10.55	(2.32)	6.65	(3.12)
High	Text-Only	25	112.52	(10.71)	11.28	(2.19)	6.64	(3.03)
	Text & Graphics	17	114.47	(8.84)	10.59	(2.29)	6.00	(2.65)
	Graphics & Audio	20	115.95	(12.05)	10.65	(1.84)	6.10	(3.01)
	Text, Graphics & Audio	20	111.65	(8.12)	9.65	(2.76)	6.00	(2.94)

Note. WM = working memory. Time 1 = immediate recall, Time 2 = delayed recall. CAT = Cognitive Abilities Test scores.

5.2.2 Map Task Descriptive Statistics

The means and standard deviations of recall scores (immediate, delayed) and CAT (Cognitive Abilities Test) scores for cognitive load (text-only; text & graphics; graphics & audio; text, graphics & audio) in the map curriculum topic, with scores for gender, WM, and motivation presented separately (see Table 5.4.).

Table 5.4.

Means and Standard Deviations of Recall Scores (Immediate, Delayed) and CAT Scores (Cognitive Abilities Test) for Cognitive Load (Text-Only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) in the Map Curriculum Topic with Scores for Gender, WM, and Motivation presented separately

Individual Differences	Variables	N	CAT Score *M*	*(SD)*	Immediate Recall Score *M*	*(SD)*	Delayed Recall Score *M*	*(SD)*
	Text-Only	58	114.67	(8.85)	4.69	(2.39)	2.05	(1.65)
	Text & Graphics	45	112.27	(9.13)	4.62	(2.99)	2.11	(1.61)
	Graphics & Audio	47	113.79	(9.03)	6.06	(2.61)	3.47	(2.15)
	Text, Graphics & Audio	48	113.73	(9.46)	4.77	(3.06)	2.65	(2.04)
Gender								
Males	Text-Only	30	115.30	(8.83)	4.50	(2.60)	2.00	(1.49)
	Text & Graphics	25	112.80	(8.94)	4.44	(3.12)	2.32	(3.12)
	Graphics & Audio	21	111.71	(7.48)	5.67	(2.37)	3.29	(2.17)
	Text, Graphics & Audio	25	116.80	(8.28)	4.60	(2.82)	2.56	(1.92)
Female	Text-Only	28	114.00	(8.99)	4.89	(2.18)	2.11	(1.83)
	Text & Graphics	20	111.60	(9.55)	4.85	(2.87)	1.85	(1.23)
	Graphics & Audio	26	115.46	(9.95)	6.38	(2.79)	3.62	(2.16)
	Text, Graphics & Audio	23	110.39	(9.69)	4.96	(3.35)	2.74	(2.20)
WM								
Low	Text-Only	13	112.54	(9.27)	3.85	(1.99)	1.69	(2.14)
	Text & Graphics	14	108.50	(8.01)	4.14	(2.54)	2.00	(1.18)
	Graphics & Audio	5	110.60	(9.53)	5.20	(2.28)	2.80	(1.30)
	Text, Graphics & Audio	9	107.00	(9.30)	3.56	(2.40)	2.11	(1.36)
High	Text-Only	10	120.60	(8.50)	4.50	(2.59)	1.70	(1.16)
	Text & Graphics	12	116.92	(7.86)	5.92	(3.83)	2.92	(2.43)
	Graphics & Audio	6	112.83	(3.55)	6.33	(3.27)	3.00	(1.55)
	Text, Graphics & Audio	9	117.44	(6.25)	4.00	(2.24)	3.00	(2.74)
Motivation								
Low	Text-Only	38	115.37	(8.99)	4.50	(2.28)	1.87	(1.23)
	Text & Graphics	26	110.96	(7.29)	4.46	(3.14)	2.31	(1.49)
	Graphics & Audio	27	114.59	(8.60)	5.30	(2.18)	3.11	(1.63)
	Text, Graphics & Audio	26	113.58	(7.16)	4.73	(2.96)	2.50	(1.92)
High	Text-Only	20	113.35	(8.65)	5.05	(2.63)	2.40	(2.23)
	Text & Graphics	19	114.05	(11.14)	4.84	(2.83)	1.84	(1.77)
	Graphics & Audio	20	112.70	(9.71)	7.10	(2.83)	3.95	(2.67)
	Text, Graphics & Audio	22	113.91	(11.80)	4.82	(3.25)	2.82	(2.20)

Note. WM = working memory. Time 1 = immediate recall, Time 2 = delayed recall. CAT = Cognitive Abilities Test scores.

5.2.3 Poem Task Descriptive Statistics

The means and standard deviations of recall scores (immediate, delayed) and CAT (Cognitive Abilities Test) scores for cognitive load (Text-Only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) in the poem curriculum topic, with scores for gender, WM, and motivation presented separately (see Table 5.5.).

Table 5.5.

Means and Standard Deviations of Recall Scores (Immediate, Delayed) and CAT Scores (Cognitive Abilities Test) for Cognitive Load (Text-Only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) in the Poem Curriculum Topic with Scores for Gender, WM, and Motivation presented separately

Ind.	Cognitive Load		CAT Score		Time 1		Time 2	
Diff.	Variables	N	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
	Text-Only	58	112.50	(8.50)	9.36	(1.65)	8.72	(2.16)
	Text & Graphics	56	113.80	(9.30)	9.55	(1.51)	9.18	(1.87)
	Graphics & Audio	46	112.26	(9.40)	8.96	(1.81)	8.59	(1.81)
	Text, Graphics & Audio	48	112.42	(8.15)	8.63	(1.58)	7.92	(2.25)
Gender								
Males	Text-Only	25	112.20	(7.52)	9.24	(1.64)	8.24	(2.54)
	Text & Graphics	25	115.96	(8.24)	9.64	(1.50)	9.21	(2.00)
	Graphics & Audio	25	113.96	(9.04)	8.88	(1.79)	8.32	(1.93)
	Text, Graphics & Audio	30	113.27	(8.26)	8.57	(1.74)	8.13	(2.00)
Female	Text-Only	33	112.73	(1.68)	9.45	(1.68)	9.09	(1.86)
	Text & Graphics	31	112.06	(9.85)	9.48	(1.55)	9.16	(1.79)
	Graphics & Audio	21	110.24	(9.63)	9.05	(1.88)	8.90	(1.64)
	Text, Graphics & Audio	18	111.00	(8.00)	8.72	(1.32)	7.56	(2.64)
WM								
Low	Text-Only	7	107.29	(11.72)	8.57	(1.13)	6.86	(3.19)
	Text & Graphics	15	106.60	(7.15)	8.53	(1.41)	8.73	(2.49)
	Graphics & Audio	14	114.14	(9.13)	8.29	(1.86)	7.64	(1.91)
	Text, Graphics & Audio	16	107.63	(6.60)	8.88	(1.71)	7.88	(1.93)
High	Text-Only	9	112.44	(9.22)	9.33	(1.58)	8.67	(1.94)
	Text & Graphics	16	118.44	(7.25)	10.00	(1.75)	9.89	(1.48)
	Graphics & Audio	7	117.29	(8.26)	8.71	(2.06)	8.71	(1.70)
	Text, Graphics & Audio	8	118.25	(7.32)	8.75	(1.58)	8.25	(2.19)
Motivation								
Low	Text-Only	34	113.29	(8.20)	9.18	(1.66)	8.53	(1.76)
	Text & Graphics	32	116.00	(9.07)	9.41	(1.48)	9.07	(1.91)
	Graphics & Audio	26	111.54	(9.66)	8.81	(1.83)	8.46	(1.84)
	Text, Graphics & Audio	29	111.41	(5.76)	8.66	(1.65)	7.45	(2.06)
High	Text-Only	24	111.38	(8.96)	9.63	(1.64)	9.00	(2.64)
	Text & Graphics	24	110.88	(8.96)	9.75	(1.57)	9.33	(1.83)
	Graphics & Audio	20	113.20	(9.21)	9.15	(1.18)	8.75	(1.80)
	Text, Graphics & Audio	19	113.95	(10.86)	8.58	(1.50)	8.63	(2.39)

Note. WM = working memory. Time 1 = immediate recall, Time 2 = delayed recall. CAT = Cognitive Abilities Test scores.

5.3 Group Comparisons

Inferential statistical analyses were carried out to investigate the effect of cognitive load (Text-Only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio), individual differences (gender, motivation, WM), and recall (immediate, delayed) were undertaken to determine an answer to the research questions “Does cognitive load (Text-Only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) affect learning outcome?" and "Do individual differences (gender, WM, motivation) moderate the effect of cognitive load on learning outcome?" Inferential statistical analyses (ANOVA, ANCOVA) were reported to determine the extent to which the IVs accounted for unique variance in recall. A summary for levels of significance for each curriculum topic can be viewed in Table 5.23.

5.3.1 Does Cognitive Load have an Impact on Learning?

The main effects and interactions were explored for cognitive load (text-only, text & graphics; graphics & audio; text, graphics & audio) across three separate curriculum topics (heart, map, poem) in two testing times (immediate recall, delayed recall). Statistical analysis was conducted to assess the impact of cognitive load on time of recall–separate analysis was conducted for each curriculum topic. To determine participants existing knowledge, a prior knowledge assessment was completed in the heart and poem curriculum topics. ANCOVA analysis was performed on curriculum topics that tested for prior knowledge (heart, poem), to determine whether within-group error variances existed, and ANOVA analysis when no prior knowledge existed (map). The analysis was factorial, i.e., included the interactions between different independent variables, rather than univariate. The means and standard deviations of these variables are presented in separate tables (Table 5.3., heart; Table 5.4., map; Table 5.5., poem). In instances where data violated the assumptions on which parametric tests are based, the non-parametric Wilcoxon signed-rank, Mann-Whitney, and independent samples Kruskal-Wallis tests were conducted. When parametric and non-parametric results concurred, parametric results were reported. When parametric and non-parametric results did not concur, the non-parametric results were reported. The between participant variables were cognitive load (text-only, text & graphics; graphics & audio; text, graphics & audio) and individual differences (gender, WM, motivation), the within participant variable was time of recall (immediate, delayed), and the covariate for ANCOVA analysis was prior knowledge.

5.3.1.1 Heart curriculum topic

The heart curriculum topic data were analysed using analysis of covariance (ANCOVA). The number of participants (male, female), mean recall scores, mean age, and mean CAT scores (with their respective standard deviations), can be viewed in Table 5.1., Table 5.2., and Table 5.3., respectively. In certain instances data violated the assumptions of normality, so both parametric and non-parametric tests were performed on the data. Non-parametric analysis concurred with parametric analysis, so parametric results were reported.

5.3.1.1.1 Analysis of covariance for the heart curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 5.6. The means and standard deviations are presented separately in Table 5.3.

Table 5.6.

ANCOVA Main Effects and Interactions for Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) in the Heart Curriculum Topic in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	292.98	1, 198	4.46	65.64 ***
Time * Prior Knowledge	2.23	1, 198	4.46	0.50
Time * Cognitive Load	12.19	3, 198	4.46	2.73 *
Prior Knowledge	377.03	1, 198	7.77	48.53 ***
Cognitive Load	13.33	3, 198	7.77	0.17

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.29. **p* < .05. ***p* < .01. ****p* < .001.

ANCOVA analysis determined there were significant main effects of time of testing, $F_{(1, 198)} = 65.64$, $MSE = 4.46$, $p < .001$, confirming students retained more information in the short term (immediate testing time), $M = 10.24$, $SE = .16$, than in the longer term (delayed testing time), $M = 5.63$, $SE = .20$. The covariate prior knowledge was significant, $F_{(1, 198)} = 48.53$, $MSE = 7.77$, $p < .001$, confirming prior knowledge influenced learning. To gain a fuller understanding of the nature of the covariate effect, additional correlation analysis was carried out, which determined that the direction of the correlation was positive, indicating that as prior knowledge scores scores increased, so did recall, $r = .43$, $p < .01$. The interaction between time and cognitive load was significant, $F_{(3, 198)} = 2.73$, $MSE = 4.46$, $p = .045$, which suggested the effect of time was not consistent across the different cognitive loads and the effect of cognitive load was not consistent across the immediate recall and delayed recall testing times.

Given the significant interaction between testing time and cognitive load, the simple effect of cognitive load was next investigated at each time of testing through univariate analysis of variance. There was no significant effect of cognitive load in immediate recall, $F_{(3, 198)} = 0.52$, $MSE = 4.80$, $p > .05$, but there was a significant effect of cognitive load in delayed recall, $F_{(3, 198)} = 3.10$, $MSE = 7.43$, $p = .028$. Bonferroni's adjusted means test investigated the pairwise differences between the cognitive load groups, while taking into account Type 1 errors. Pairwise differences between the cognitive load conditions in delayed recall were investigated with prior knowledge as a covariate, and results revealed a significant difference between conditions one (text-only) and four (text, graphics and audio), $MD = 1.50$, $p = .020$, confirming recall from the text, graphics and audio group, $M = 6.56$, $SE = 0.38$, was significantly better than the text-only group, M = 5.06, SE = 0.33. All other pairwise comparisons failed to reach significance. Results confirmed, in the delayed testing time, that text, graphics and audio was the better condition for the retention of information in the heart topic, relative to text-only.

5.3.1.1.2 Heart curriculum topic summary

In answer to the research question, "Does cognitive load affect learning outcome?" investigation revealed cognitive load appeared to modulate the effects of recall in testing time in the science curriculum topic. Analysis confirmed there was no significant difference in learning outcome between students in different cognitive load conditions in the short term, i.e., within an hour after being presented with the testing material. However, students retained significantly more information in the text, graphics and audio condition after a longer period of time i.e., four to six weeks later, relative to the text-only condition. The cognitive ability of participants in each group was similar, which confirmed the difference found was not determined by the fact that there was a difference in CAT scores between groups (see Table 5.1). However, unequal sample sizes, which ranged from 35-68 participants, may have impacted on the results. Therefore, any inference in causality must be made with caution.

5.3.1.2 Map curriculum topic

The map curriculum topic data were analysed using analysis of variance (ANOVA). In certain instances data violated the assumptions on which parametric tests are based. In delayed recall Levene's test of equality of error variances was significant, $p = .027$, so both parametric and non-parametric tests were performed on the data. Results concurred, so parametric results were reported.

5.3.1.2.1 Analysis of variance for the map curriculum topic

Results of the factorial ANOVA analysis are summarised in Table 5.7. Means and standard deviations are presented separately in Table 5.4.

Table 5.7.

ANOVA Main Effects and Interactions in the Map Topic for Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) in Two Testing Times (Immediate, Delayed)

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	597.09	1, 194	2.93	203.72 ***
Time * Cognitive Load	1.36	3, 194	2.93	0.46
Cognitive Load	42.23	3, 194	8.14	5.19 **

*p < .05. **p < .01. ***p < .001.

ANOVA analysis determined there were significant main effects of time of testing, $F_{(1, 194)} = 206.23$, $MSE = 3.04$, $p < .001$, confirming participants recalled more information in the immediate recall testing time, $M = 5.04$, $SE = .20$, than in the delayed recall testing time, $M = 2.51$, $SE = .13$. There were significant main effects of cognitive load, $F_{(3, 194)} = 5.19$, $MSE = 8.14$, $p = .002$. As time was not a significant factor in the interaction, the data were collapsed across the two time points, taking the mean of immediate recall and delayed recall scores. Bonferroni's adjusted means test investigated the pairwise differences between the cognitive load groups, while taking into account Type 1 errors. Results revealed a significant difference between conditions one (text-only) and three (graphics and audio), $MD = -1.43$, $p = .001$, and two (text and graphics) and three (graphics and audio), $MD = -1.49$, $p = .002$, confirming recall from the graphics and audio sample group, $M = 4.77$, $SE = 0.28$, was significantly better than the text-only, M = 3.34, SE = 0.26, and text and graphics, M = 3.28, SE = 0.29, sample groups. All other pairwise comparisons failed to reach significance. Results confirmed, in the delayed testing time, that graphics and audio was better for retention, relative to text-only and text and graphics.

5.3.1.2.2 Map curriculum topic summary

In answer to the research question, "Does cognitive load affect learning outcome?" investigation revealed cognitive load appeared to modulate the effects of recall in the map curriculum topic. Analysis confirmed students recalled more information from the graphics and audio condition relative to text-only and text and graphics conditions. The cognitive ability of students in each group was similar, which demonstrated the difference found was not determined by the fact that there was a difference in CAT scores between cognitive load groups (see Table 5.1.). However, unequal sample sizes, which ranged from 45-58 students, may have had an impact on the results (see Table 5.4.). Therefore, any inferences regarding causality must be made with caution.

5.3.1.3 Poem curriculum topic

The poem curriculum topic data were analysed using analysis of covariance (ANCOVA). The number of participants (male, female), mean recall scores, mean age, and mean CAT scores (with their respective standard deviations), can be viewed in Table 5.1., Table 5.2., and Table 5.5., respectively. In certain instances data violated the assumptions on which parametric tests are based, so both parametric and non-parametric tests were performed on the data. Results concurred, so parametric results were reported.

5.3.1.3.1 Analysis of covariance for the poem curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 5.8. Means and standard deviations are presented separately in Table 5.5.

Table 5.8.

ANCOVA Main Effects and Interactions in the Poem Topic for Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) in Two Testing Times (Immediate, Delayed) with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	32.74	1, 203	1.91	17.12 ***
Time * Prior Knowledge	6.07	1, 203	1.91	3.17
Time * Cognitive Load	0.86	3, 203	1.91	0.45
Prior Knowledge	12.96	1, 203	4.83	2.68
Cognitive Load	22.51	3, 203	4.83	4.66 **

Note. Covariates appearing in the model are evaluated at the following values: Poem Prior Knowledge Raw Score /6 = .23

ANCOVA analysis determined there was a significant main effect of time of testing, $F_{(1, 203)} = 17.12$, $MSE = 1.91$, $p < .001$, confirming students recalled more information in the short term (immediate testing time), $M = 9.12$, $SE = .11$, than in the longer term (delayed testing time), $M = 8.60$, $SE = .14$. There was also a significant main effect of cognitive load, $F_{(3, 203)} = 4.66$, $MSE = 4.83$, $p = .004$. As the interaction between testing time and cognitive load was not significant, data obtained at both testing times was collapsed and the average of scores used for further analysis investigating the main effect of cognitive load. Bonferroni's adjusted means test investigated the pairwise differences between the cognitive load groups, while taking into account Type 1 errors. Results revealed a significant difference between conditions two (text and graphics) and four (text, graphics and audio), *MD* = 1.08, *p* = .003, confirming recall from the text and graphics sample group, *M* = 9.37, *SE* = 0.21, was significantly better than the text, graphics and audio, *M* = 8.29, *SE* = 0.22, sample group. All other pairwise comparisons failed to reach significance. Results confirmed that text and graphics was better for retention, relative to text, graphics and audio.

5.3.1.3.2 Poem curriculum topic summary

In answer to the research question, "Does cognitive load affect learning outcome?" investigation revealed cognitive load appeared to modulate the effects of recall in the poem curriculum topic. Analysis suggested text, graphics and audio was a less successful combination for learning for students, relative to text and graphics. The cognitive ability of students in each group was similar, which demonstrated the difference found was not determined by the fact that there was a difference in CAT scores between cognitive load groups (see Table 5.5.). However, unequal sample sizes, which ranged from 46-58 students, may have had an impact on the results. Therefore, any inferences regarding causality must be made with caution.

5.3.2 Does Gender Moderate the Effect of Cognitive Load?

Main effects and interactions were explored for gender (males, females) and cognitive load conditions (text-only; text & graphics; graphics & audio; text, graphics & audio) across three separate curriculum topics (heart, map, poem) in two testing times (immediate, delayed) to determine an answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?" The between participant variables were cognitive load and gender, the within participant variable was time of recall, and the covariate for the ANCOVA analysis was prior knowledge.

5.3.2.1 Heart curriculum topic

The heart curriculum topic data were analysed using analysis of covariance (ANCOVA). The number of participants, as well as CAT mean and standard deviation of scores for each condition, can be viewed in Table 5.3. In the immediate testing time, data violated the assumptions on which parametric tests are based. Levene's test of equality of error variances was significant, $p = .02$, so both parametric and non-parametric tests were performed on the data. Results concurred, and parametric results were reported.

5.3.2.1.1 Analysis of covariance for the heart curriculum topic

Results of the ANCOVA analysis are summarised in Table 5.9. Means and standard deviations are presented in Table 5.3.

Table 5.9.

ANCOVA Main Effects and Interactions in the Heart Task for Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) and Gender (Male, Female) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	261.79	1, 194	4.44	58.91 ***
Time * Prior Knowledge	4.48	1, 194	4.44	1.01
Time * Cognitive Load	13.30	3, 194	4.44	2.99 *
Time * Gender	7.58	1, 194	4.44	1.71
Prior Knowledge	386.18	1, 194	7.71	50.08 ***
Cognitive Load	12.89	3, 194	7.71	1.67
Gender	18.06	1, 194	7.71	2.34
Cognitive Load * Gender	5.74	3, 193	7.71	0.74
Time * Cognitive Load * Gender	4.51	3, 194	4.44	1.02

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.29. **p* < .05. ***p* < .01. ****p* < .001.

ANCOVA analysis determined there was a significant effect of time of testing, $F_{(1, 194)} = 58.91$, $p < .001$, confirming that students recalled more information in the short term (immediate testing time), $M = 10.24$, $SE = .16$, than in the longer term (delayed testing time), $M = 5.64$, $SE = .20$. The covariate prior knowledge was significant, $F_{(1, 194)} = 50.08$, $p < .01$, which indicated prior knowledge had an influence on learning. To gain a fuller understanding of the nature of the covariate effect, additional correlation analysis was carried out, which determined that the direction of the correlation was positive, indicating that as prior knowledge scores scores increased, so did recall, $r = .43$, $p < .01$. The interaction between testing time and cognitive load was significant, and these findings were previously identified and discussed. Given the lack of significant main effect or interactions involving gender, no further analysis was necessary.

5.3.2.1.2 Heart curriculum topic summary

In answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?" investigation revealed gender did not appear to modulate the effects of either cognitive load or testing time in the science curriculum topic. Analysis confirmed, in line with CAT scores across groups (see Table 5.3), that there was no difference to learning outcome in the heart topic between boys and girls in each of the cognitive load conditions within an hour after being presented with the testing material as well as four to six weeks later.

5.3.2.2 Map curriculum topic

The map curriculum topic data were analysed using analysis of variance (ANOVA). In certain instances data violated the assumptions on which parametric tests are based. Levene's test of equality of error variances was significant in delayed recall, p = .007, so both parametric and non-parametric tests were performed on the data. Results did not concur, so non-parametric Wilcoxon signed-rank, Mann-Whitney, and Kruskal-Wallis test results were reported. The number of participants, as well as CAT mean and standard deviation of scores for each condition, can be viewed in Table 5.4.

5.3.2.2.1 Non-parametric analysis for the map curriculum topic

Non-parametric data analysis identified levels of significance between time, gender, and cognitive load (see Table 5.10.).

Table 5.10.

Non-Parametric Analyses in the Map Task, Identifying Levels of Significance for Interactions between Time (Immediate, Delayed), Gender (Males, Females) and Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio)

Interactions	Measures					
	Median Time	*Median Ind. Diff.*	*U*	*T*	*Z*	*r*
Time x Gender Male	4.00, 2.00			212.00		-0.72 ***
Time x Gender Female	5.00, 2.00			110.00		-0.78 ***
Time 1 x Gender x Text-Only		4.00, 5.00	480.50		0.95	0.13
Time 1 x Gender x Text & Graphics		4.00, 4.50	278.00		0.65	0.10
Time 1 x Gender x Graphics & Audio		5.00, 5.50	307.50		0.75	0.11
Time 1 x Gender x Text, Graphics & Audio		4.00, 4.00	296.00		0.18	0.03
Time 2 x Gender x Text-Only		2.00, 2.00	422.50		0.04	0.01
Time 2 x Gender x Text & Graphics		2.00, 2.00	241.00		-0.21	-0.03
Time 2 x Gender x Graphics & Audio		4.00, 3.00	286.00		0.28	0.04
Time 2 x Gender x Text, Graphics & Audio		2.00, 2.00	296.00		0.18	0.03

Note. Time 1 = immediate recall, Time 2 = delayed recall. *Median (Time)* = immediate recall, delayed recall. *Median (Ind. Diff.)* = individual difference: gender. *U* = Mann-Whitney U. *T* = Wilcoxon signed-rank.
**p* < .05. ***p* < .01. ****p* < .001.

Wilcoxon signed-rank tests identified that performance was significantly higher in the immediate recall testing time than in the delayed recall testing time for both boys and girls (see Table 5.11.). Mann-Whitney analysis identified no significant differences between boys and girl in any of the cognitive load conditions in both the immediate testing time and delayed testing time. Kruskal-Wallis analysis found a significant difference within cognitive load conditions for girls in delayed testing time, $H_{(3)} = 10.43$, $p = .015$. Pairwise comparisons, with adjusted *p*-values, identified a significant difference for girls between: text-only and graphics & audio; and text & graphics and text, graphics & audio, where girls retained more information in the graphics and audio condition in the delayed testing time relative to the text-only and text & graphics conditions (see Table 5.11). No other pairwise comparisons were significant.

Table 5.11.

Kruskal-Wallis Adjusted p-Value Pairwise Comparisons for Females and Cognitive Load Conditions in the Map Topic in the Delayed Testing Time

Gender (Females) Delayed Testing Time	*r*	Text-Only	Text & Graphics	Graphics & Audio	Text, Graphics & Audio
		Mean Rank	*Mean Rank*	*Mean Rank*	*Mean Rank*
Females	-0.38	41.59		62.69 *	
Females	-0.39		40.70	62.69 *	

Note. *r* = effect size. **p* < .05. ***p* < .01.

5.5.2.2.2 Map curriculum topic summary

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?” investigation revealed gender appeared to modulate the effects of cognitive load and testing time in the geography curriculum topic. No significant differences were found between boys and girls, but analysis confirmed girls retained more information in the graphics and audio condition relative to the text-only and text & graphics conditions four to six weeks after being presented with the testing material. Female CAT scores were the highest in the graphics and audio condition (but only just when compared to text-only), and this may have impacted on performance (see Table 5.4.). There was no significant difference to learning outcome for boys in all cognitive load conditions in both testing times in the map topic. However, unequal sample sizes, which ranged from 20-30 participants, may have had a impact on the results. Therefore, any inferences regarding causality must be made with caution.

5.3.2.3 The English poem curriculum topic

The poem curriculum topic data were analysed using analysis of covariance (ANCOVA). The number of participants, mean scores and standard deviations, and CAT mean scores for each sample group, can be viewed in Table 5.5.

5.3.2.3.1 Analysis of covariance for the poem curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 5.12. Means and standard deviations are presented separately in Table 5.5.

Table 5.12.

ANCOVA Main Effects and Interactions in the Poem Task for Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) and Gender (Male, Female) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	34.77	1, 199	1.92	18.12 ***
Time * Prior Knowledge	5.26	1, 199	1.92	2.74
Time * Cognitive Load	1.24	3, 199	1.92	0.64
Time * Gender	0.45	1, 199	1.92	0.24
Prior Knowledge	12.45	1, 199	4.87	2.56
Cognitive Load	21.95	3, 199	4.87	4.51 **
Gender	2.90	1, 199	4.87	0.60
Cognitive Load * Gender	2.96	3, 199	4.87	0.61
Time * Cognitive Load * Gender	1.83	3, 199	1.92	0.95

Note. Covariates appearing in the model are evaluated at the following values: Poem Prior Knowledge Raw Score /6 = 0.23.
*p < .05. **p < .01. ***p < .001.

ANCOVA analysis determined there were significant main effects of time of testing and cognitive load, which have been previously identified. There were no significant interactions. Given the lack of significant interactions involving gender, no further analysis was necessary.

5.3.2.3.2 Poem curriculum topic summary

In answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?" investigation revealed gender did not appear to modulate the effects of either cognitive load or testing time in the English curriculum topic. Analysis confirmed there was no significant difference in the retention of information in the poem topic between boys and girls in each of the cognitive load conditions (text-only; text & graphics; graphics & audio; text, graphics & audio) in both testing times (immediate testing time, delayed testing time) despite slight variations in CAT scores, (see Table 5.5.).

5.3.3 Does Working Memory Moderate the Effect of Cognitive Load?

Main effects and interactions were explored for WM and cognitive load conditions across three separate curriculum topics in two testing times to determine an answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?" To determine high and low WM, scores above the median were classified as high WM, and scores equal to or lower than the median were classified as low WM. As participation for the WM assessment was voluntary, data collection resulted in small, uneven sample sizes. The between participant variables were cognitive load and WM, the within participant variable was time of recall, and the covariate for the ANCOVA analysis was prior knowledge.

5.3.3.1 Heart curriculum topic

The heart curriculum topic data were analysed using analysis of covariance (ANCOVA). The sample included 94 participants aged 11-14 years (*M* = 12 years 6 months, *SD* = 0.85). The number of participants, as well as CAT mean and standard deviation of scores for each condition, can be viewed in Table 5.3. In certain instances data violated the assumptions of normality, so both parametric and non-parametric tests were performed on the data. Non-parametric test results concurred with the parametric analysis, so parametric results were reported.

5.3.3.1.1 Analysis of covariance for the heart curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 5.13. Means and standard deviations are presented separately in Table 5.3.

Table 5.13.

ANCOVA Main Effects and Interactions in the Heart Topic for Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) and WM (High, Low) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	131.75	1, 85	4.90	26.87 ***
Time * Prior Knowledge	0.44	1, 85	4.90	0.09
Time * Cognitive Load	12.39	3, 85	4.90	2.53
Time * WM	0.01	1, 85	4.90	0.00
Prior Knowledge	203.92	1, 85	6.75	30.22 ***
Cognitive Load	4.76	3, 85	6.75	0.71
WM	4.37	1, 85	6.75	4.37 *
Cognitive Load * WM	11.51	3, 85	6.75	1.71
Time * Cognitive Load * WM	4.19	3, 85	4.90	0.86

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.30. **p* < .05. ***p* < .01. ****p* < .001.

ANCOVA analysis determined there was a significant main effect of time of testing, $F_{(1, 85)} = 26.88$, $p < .01$, confirming that students recalled more information in the short term (immediate testing time), *M* = 9.84, *SE* = .23, than in the longer term (delayed testing time), *M* = 5.40, *SE* = .29. The covariate prior knowledge was significant, $F_{(1, 85)}$ =

30.22, $p < .01$, which indicated prior knowledge had an influence on learning. To gain a fuller understanding of the nature of the covariate effect, additional correlation analysis was carried out, which determined that the direction of the correlation was positive, indicating that as prior knowledge scores scores increased, so did recall, $r = .50$, $p < .01$. The main effect of WM was significant, $F_{(1, 85)} = 4.37$, $MSE = 6.75$, $p = .040$, confirming high WM participants recalled more information, $M = 8.03$, $SE = .30$, than low WM participants, $M = 7.20$, $SE = .27$. There were no significant interactions. Given the lack of significant interactions involving WM, no further analysis was necessary.

5.3.3.1.2 Heart curriculum topic summary

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?” investigation revealed WM did not modulate the effects of cognitive load in the heart curriculum topic. Analysis confirmed, in certain instances, that prior knowledge influenced learning, and that high WM participants outperformed low WM participants. However, small and unequal sample sizes, which ranged from 6-17 participants, as well as varied CAT scores across groups (see Table 5.3.), may have had an impact on the results. Therefore, any inferences regarding causality must be made with caution.

5.3.3.2 Map curriculum topic

The map curriculum topic data were analysed using analysis of variance (ANOVA), in which the sample included 78 participants aged 11-14 years (M = 12 years 7 months, $SD = 0.85$). In certain instances data violated the assumptions of normality, so both parametric and non-parametric tests were performed on the data. Non-parametric test results concurred with parametric analysis, so parametric results were reported. The number of participants, as well as CAT mean and standard deviation of scores for each condition, can be viewed in Table 5.4.

4.3.3.2.1 Analysis of variance for the map curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 5.14. Means and standard deviations are presented separately in Table 5.4.

Table 5.14.

ANCOVA Main Effects and Interactions in the Map Task for Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) and WM (High, Low) in Two Testing Times (Immediate, Delayed)

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	177.16	1, 70	2.36	74.94 ***
Time * Cognitive Load	4.25	3, 70	2.36	1.80
Time * WM	2.64	1, 70	2.36	1.12
Cognitive Load	12.69	3, 70	8.74	1.45
WM	18.30	1, 70	8.74	2.10
Cognitive Load * WM	1.72	3, 70	8.74	0.20
Time * Cognitive Load * WM	1.11	3, 70	2.36	0.47

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.30.
*p < .05. **p < .01. ***p < .001.

ANOVA analysis determined there was a significant effect of time of testing, $F_{(1, 70)} = 74.74$, $p < .001$, confirming that students recalled more information about the map within an hour of being presented with the learning material (immediate testing time), M = 4.69, SE = .33, than after four to six weeks (delayed testing time), M = 2.40, SE = .22. There was no significant main effect of cognitive load or WM (see Table 5.14.), and no significant interactions. Given the lack of significant interactions involving WM, no further analysis was necessary.

5.3.3.2.2 Map curriculum topic summary

In answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of cognitive load on learning outcome?" investigation revealed WM did not appear to modulate the effects of either cognitive load or testing time in the geography curriculum topic. Analysis confirmed there was no significant difference in learning outcome between students with high or low WM in each of the cognitive load conditions (text-only; text & graphics; graphics & audio; text, graphics & audio) within an hour of being presented with the material and after four to six weeks.

5.3.3.3 Poem curriculum topic

The poem curriculum topic data were analysed using analysis of covariance (ANCOVA). The sample included 92 participants aged 11-14 years (M = 12 years 8 months, SD = 0.81). The number of participants (male, female), as well as recall score means, age means, and CAT means (with their respective standard deviations), can be viewed in Table 5.1., Table 5.2., and Table 5.4., respectively.

5.3.3.3.1 Analysis of covariance for the poem curriculum topic

Results of the factorial ANOVA analysis are summarised in Table 5.15. Means and standard deviations are presented separately in Table 5.5.

Table 5.15.

ANCOVA Main Effects and Interactions in the Poem Task for Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) and WM (High, Low) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	15.65	1, 83	1.97	7.93 **
Time * Prior Knowledge	8.21	1, 83	1.97	4.16 *
Time * Cognitive Load	2.75	3, 83	1.97	1.39
Time * WM	1.03	1, 83	1.97	0.52
Prior Knowledge	0.47	1, 83	5.12	0.09
Cognitive Load	10.39	3, 83	5.12	2.03
WM	28.71	1, 83	5.12	5.61 *
Cognitive Load * WM	3.16	3, 83	5.12	0.62
Time * Cognitive Load * WM	1.39	3, 83	1.97	0.70

Note. Covariates appearing in the model are evaluated at the following values: Poem Prior Knowledge Raw Score /6 = 0.15. *$p < .05$. **$p < .01$. ***$p < .001$.

ANCOVA analysis determined there was a significant main effect of time of testing, $F_{(1, 83)} = 7.93$, $p < .01$, confirming that students recalled more information about the poem within an hour of being presented with the learning material (immediate testing time), $M = 8.88$, $SE = .19$, than after four to six weeks (delayed testing time), $M = 8.34$, $SE = .23$. To gain a fuller understanding of the nature of the covariate effect, additional correlation analysis was carried out, which determined that the direction of the correlation was positive, indicating that as prior knowledge scores scores increased in the delayed testing time, so did recall, $r = .14$, $p < .05$. The main effect of WM was significant, $F_{(1, 83)} = 5.61$, $MSE = 5.12$, $p = .020$, confirming high WM students recalled more information, $M = 9.03$, $SE = .27$, than low WM students, $M = 8.18$, $SE = .24$. The interaction between time and prior knowledge was significant, $F_{(1, 83)} = 4.16$, $MSE = 1.97$ $p = .045$, which confirmed prior knowledge influenced learning outcome. There were no other significant interactions. Given the lack of significant interactions involving WM, no further analysis was necessary.

5.3.3.3.2 Poem curriculum topic summary

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of cognitive load on learning outcome?” investigation revealed WM modulated the effect of recall in the poem curriculum topic. Analysis confirmed high WM students performed significantly better than low WM students. Prior knowledge also had an influence on learning outcome and had both a positive and negative affect on students’ recall scores. However, small and unequal sample sizes, which ranged from 7-16 participants, could have influenced the results. Therefore, any inferences regarding causality must be made with caution.

5.3.4 Does Motivation Moderate the Effect of Cognitive Load?

Main effects and interactions were explored for motivation and cognitive load conditions across three separate curriculum topics in two testing times to determine an answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?” To determine high and low motivation, scores above the median were classified as high motivation, and scores equal to or lower than the median were classified as low motivation. The between participant variables were cognitive load and motivation, the within participant variable was time of recall, and the covariate for the ANCOVA analysis was prior knowledge.

5.3.4.1 Heart curriculum topic

The heart curriculum topic data were analysed using analysis of covariance (ANCOVA). In immediate recall, data violated the assumptions on which parametric tests are based. Levene’s test of equality of error variances was significant, $p = .004$, so both parametric and non-parametric tests were performed on the data. Results concurred, so parametric findings were reported. The number of participants, as well as CAT mean and standard deviation of scores for each condition, can be viewed in Table 5.3.

5.3.4.1.1 Analysis of covariance for the heart curriculum topic

Results of the factorial ANCOVA analysis are summarised in Table 5.16. Means and standard deviations are presented separately in Table 5.3.

Table 5.16.

ANCOVA Main Effects and Interactions in the Heart Task for Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) and Motivation (High, Low) in Two Testing Times (Immediate, Delayed), with Prior Knowledge as a Covariate

Source	*Mean Square*	*df*	*MSerror*	*F*
Time	259.01	1, 194	4.48	57.79 ***
Time * Prior Knowledge	3.98	1, 194	4.48	0.89
Time * Cognitive Load	9.21	3, 194	4.48	2.06
Time * Motivation	4.74	1, 194	4.48	1.06
Prior Knowledge	322.53	1, 194	7.38	43.73 ***
Cognitive Load	7.18	3, 194	7.38	0.97
Motivation	28.40	1, 194	7.38	3.85
Cognitive Load * Motivation	24.98	3, 194	7.38	3.39 *
Time * Cognitive Load * Motivation	2.33	3, 194	4.48	0.52

Note. Covariates appearing in the model are evaluated at the following values: Heart Prior Knowledge Raw Score /9 = 3.29.
**p* < .05. ***p* < .01. ****p* < .001.

ANCOVA analysis determined there was a significant effect of time of testing $F_{(1, 194)} = 57.79$, $p < .001$, confirming that students recalled more information about the heart within an hour of being presented with the learning material (immediate testing time), $M = 10.26$, $SE = .16$, than after four to six weeks (delayed testing time), $M = 5.70$, $SE = .20$. There was a significant main effect of prior knowledge, $F_{(1, 194)} = 43.73$, $p < .001$, confirming prior knowledge influenced learning. To gain a fuller understanding of the nature of the covariate effect, additional correlation analysis was carried out, which determined that the direction of the correlation was positive, indicating that as prior knowledge scores scores increased, so did recall, $r = .43$, $p < .01$.

The interaction between cognitive load and motivation was significant, $F_{(3, 194)} = 3.39$, $MSE = 7.38$, $p = .019$, which confirmed the effect of cognitive load was not consistent for high and low motivation sample groups, and the effect of motivation was not consistent for all cognitive load conditions. Given the significant interaction between cognitive load and motivation, further simple effect tests were conducted. As time was not a significant factor in the interaction, the data were collapsed across the two time points, taking the mean of immediate recall and delayed recall scores. For this analysis Levene's test of equality of error variances was not significant, $p = 263$. Univariate analysis of variance tests were conducted separately on data from high and low motivation groups. There was a significant simple effect of cognitive load in the low motivation group, $F_{(3, 116)} = 4.56$, $MSE = 3.67$, $p = .005$, but not in the high motivation group, $F_{(3, 77)} = 0.66$, $MSE = 3.76$, $p = 577$. To further examine the significant simple effect of cognitive load in the low motivation group, Bonferroni's adjusted means test investigated the pairwise differences between the cognitive load groups, while taking into account Type 1 errors. Results revealed a significant difference between conditions one (text-only) and four (graphics and audio), *MD* = -1.56, *p* = .005, confirming recall from the text, graphics and audio sample group, *M* = 8.70, *SE* = 0.35, was significantly better than the text-only, M = 7.14, SE = 0.29, sample group. All other pairwise comparisons failed to reach significance. Results confirmed that for the low motivation sample, text, graphics and audio was better for retention, relative to text-only.

5.3.4.1.2 Heart curriculum topic summary

In answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?" investigation revealed motivation appeared to modulate the effects of cognitive load in the science curriculum topic. Analysis confirmed high motivation students performed better than low motivation students. Low motivation students recalled more information in the text, graphics and audio condition relative to the text-only condition. There were no significant differences in retention for high motivation participants in the heart topic in each of four cognitive load conditions (text-only; text & graphics; graphics & audio; text, graphics & audio). However, the CAT mean score for low motivation participants was slightly higher in the text, graphics and audio condition, and this, together with the variation in sample sizes which ranged from 17-43 participants, could have influenced the results (see Table 5.18.). Therefore, any inferences regarding causality must be made with caution.

5.3.4.2 Map curriculum topic

The map curriculum topic data were analysed using analysis of variance (ANOVA). In several instances, data violated the assumptions on which parametric tests are based. Levene's analysis of equality of error variances in delayed recall was significant, *p* = .001, so both parametric and non-parametric tests were performed on the data. Results differed, so non-parametric Wilcoxon signed-rank and independent samples Kruskal-Wallis tests were reported. The number of participants, as well as CAT mean and standard deviation of scores for each condition, can be viewed in Table 5.4.

5.3.4.2.1 Non-parametric analysis for the map curriculum topic

Results of the non-parametric analyses are summarised in Table 5.17.

Table 5.17.

Non-Parametric Analysis of the Map Task Identifying Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) and Motivation (High, Low) in Two Testing Times (Immediate, Delayed)

	Measures					
Interactions	*Median (Time)*	*Median (Motivation)*	*U*	*T*	*Z*	*r*
Time x Low Mot.	4.00, 2.00			207.50		-0.77 ***
Time x High Mot.	5.00, 2.00			312.50		-0.76 ***
Time 1 x Mot. x CL1		5.00, 4.00	418.00		0.63	0.08
Time 1 x Mot. x CL2		4.00, 3.50	277.50		0.71	0.11
Time 1 x Mot. x CL3		7.00, 5.00	365.50		2.07	0.30 *
Time 1 x Mot. x CL4		4.50, 4.00	283.50		-0.52	0.07
Time 2 x Mot. x CL1		2.00, 2.00	406.50		0.45	0.06
Time 2 x Mot. x CL2		1.00, 2.00	191.00		-1.34	-0.00
Time 2 x Mot. x CL3		4.00, 3.00	316.00		1.00	0.15
Time 2 x Mot. x CL4		2.00, 2.00	319.00		0.70	0.10

Note. Time 1 = immediate recall; Time 2 = delayed recall. CL1-CL4 = cognitive load conditions (CL1 = text-only, CL2 = text & graphics, CL3 = graphics & audio, CL4 = text, graphics & audio. Ind. Diff. = individual difference (gender and motivation). *Median (Time)* = immediate recall, delayed recall. Mot. = motivation. *U* = Mann-Whitney U. *T* = Wilcoxon signed-rank. **p* < .05. ***p* < .01. ****p* < .001.

Wilcoxon signed-rank tests identified that performance was significantly higher in the immediate recall testing time than in the delayed recall testing time for both high and low motivation participants. Mann-Whitney analysis identified a significant difference between high and low motivation participants in the graphics and audio condition in the immediate recall testing time, in which high motivation participants performed better than low motivation participants (see Table 5.17.).

For the low motivation group, Kruskal-Wallis analysis identified that there was no significant difference between cognitive load conditions in immediate recall, $H_{(3)} = 3.78$, $p = .286$, but there was a significant difference in the delayed recall testing time, $H_{(3)} = 11.30$, $p = .010$. Results of significant pairwise comparisons, with adjusted *p*-values, can be viewed in Table 5.18. No other pairwise comparisons were significant. Findings suggested that students with low motivation performed better in the graphics and audio condition four to six weeks after being presented with the learning material (delayed testing time), relative to the text-only condition.

For the high motivation group, Kruskal-Wallis analysis identified that there was a significant difference between cognitive load conditions in immediate recall, $H_{(3)} = 8.77$, $p = .033$, and in delayed recall, $H_{(3)} = 8.02$, $p = .046$. Results of significant pairwise comparisons, with adjusted *p*-values, can be viewed in Table 5.18. No other pairwise comparisons were significant. Findings identified that the high motivation sample group retained more information in the graphics and audio condition in immediate recall relative to the text, graphics and audio condition, and in the graphics and audio condition in delayed recall, relative to the text and graphics condition.

Table 5.18.

Kruskal-Wallis Pairwise Comparisons, with Adjusted p-Values, for Cognitive Load Conditions in the Map Task, with Data for Low and High Motivation Groups Presented Separately

			CL1	CL2	CL3	CL4
Individual Difference	Time	*r*	*Mean Rank*	*Mean Rank*	*Mean Rank*	*Mean Rank*
Low Motivation	Time 2	-.41	49.43		76.65 **	
High Motivation	Time 1	.41			54.30 *	35.36
High Motivation	Time 2	-.43		31.61	51.78 *	

Note. Time 1 = immediate recall; Time 2 = delayed recall. CL1-4 = cognitive load conditions (CL1 = text-only, CL2 = text & graphics, CL3 = graphics & audio, CL4 = text, graphics & audio). [a] $p = .052$. *$p < .05$. **$p < .01$. ***$p < .001$.

5.3.4.2.2 Map curriculum topic summary

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?” non-parametric investigation revealed motivation appeared to modulate the effects of cognitive load and testing time in the map curriculum topic. Analysis suggested, within an hour of being presented with the learning material (immediate testing time), that high motivation students performed better than low motivation students in the graphics and audio condition, and that high motivation students performed better in the graphics and audio condition relative to the text, graphics and audio condition. After four to six weeks of being presented with the learning material (delayed testing time), high motivation students performed better in the graphics and audio condition relative to the text and graphics condition. Also, in delayed testing time, low motivation students performed better in the graphics and audio condition, relative to the text-only condition. However, unequal samples sizes, which ranged from 19-38 participants (see Table 5.5.), could have influenced the significance of results. Therefore, any inferences regarding causality must be made with caution.

5.3.4.3 Poem curriculum topic

The poem curriculum topic data were analysed using analysis of covariance (ANCOVA). In several instances, data violated the assumptions on which parametric tests are based, so both parametric and non-parametric tests were performed on the data. Results differed between the two types of tests, and so findings from Wilcoxon signed-rank and Kruskal-Wallis tests were reported. The number of participants, as well as CAT mean and standard deviation of scores for each condition, can be viewed in Table 5.5.

5.3.4.3.1 Analysis of variance for the poem curriculum topic

Results of the Wilcoxon signed-rank tests, Mann-Whitney analysis, and Kruskal-Wallis analysis are summarised in Table 5.19.

Table 5.19.

Non-Parametric Analysis of the Poem Task Identifying Cognitive Load (Text-only; Text & Graphics; Graphics & Audio; Text, Graphics & Audio) and Motivation (High, Low) in Two Testing Times (Immediate, Delayed)

	Measures					
Interactions	*Median (Time)*	*Median (Motivation)*	*U*	*T*	*Z*	*r*
Time x Low Mot.	9.00, 9.00			1326.00		-0.29 **
Time x High Mot.	10.00, 9.00			834.00		-0.17
Time 1 x Mot. x CL1		10.00, 9.00	475.50		1.08	0.14
Time 1 x Mot. x CL2		10.00, 10.00	424.50		0.69	0.09
Time 1 x Mot. x CL3		9.50, 9.00	290.00		0.68	0.10
Time 1 x Mot. x CL4		9.00, 9.00	269.00		-0.14	-0.02
Time 2 x Mot. x CL1		9.00, 9.00	478.50		1.13	0.15
Time 2 x Mot. x CL2		10.00, 9.00	421.00		0.62	0.08
Time 2 x Mot. x CL3		9.00, 8.50	290.50		0.69	0.10
Time 2 x Mot. x CL4		9.00, 8.00	375.00		2.13	0.31 *

Note. Ind. Diff. = individual difference (motivation) Time 1 = immediate recall; Time 2 = delayed recall. CL1-CL4 = cognitive load conditions (CL1 = text-only, CL2 = text & graphics, CL3 = graphics & audio, CL4 = text, graphics & audio. Mot. = motivation. *Median (Time)* = immediate recall, delayed recall. *U* = Mann-Whitney U. *T* = Wilcoxon signed-rank. *p < .05. **p < .01. ***p < .001.

Wilcoxon signed-rank tests identified that performance was significantly higher in the immediate recall testing time than in the delayed recall testing time for both high and low motivation groups (see Table 5.21.). Mann-Whitney analysis identified that there was a significant difference between low motivation and high motivation groups in the text, graphics and audio condition in the delayed recall testing time, in which the high motivation students achieved a higher mean rank score (29.74) than the low motivation students (21.07). No other significant differences were identified. Findings suggested that high motivation students performed better than low motivation students in the text, audio, and graphics condition after four to six weeks of being presented with the learning material.

Kruskal-Wallis analysis identified no significant difference between cognitive load conditions for the low motivation group in immediate recall, $H_{(3)} = 3.19$, $p = .364$. However, in delayed recall there was a significant difference between low motivation stuents, $H_{(3)} = 9.01$, $p = .029$. Pairwise comparisons, with adjusted *p*-values, identified a significant difference between text & graphics and text, graphics & audio, $p = .019$, $r = .38$, in which students in the text and graphics group achieved a higher mean rank score (71.78) than students in the text, graphics and audio group (45.64). No other pairwise comparisons were significant. Findings suggested that low motivation students performed better in the text and graphics condition in the delayed recall testing time, relative to the text, graphics and audio condition. Kruskal-Wallis analysis identified no significant difference between cognitive load conditions for high motivation groups in immediate recall, $H_{(3)} = 5.71$, $p = .127$ or delayed recall, $H_{(3)} = .812$, $p = .847$.

5.3.4.3.2 Poem curriculum topic summary

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?” motivation appeared to modulate the effects of recall. Analysis confirmed, in delayed recall, that high motivation students performed better than low motivation students in the text, graphics & audio condition, and that low motivation students performed better in the text & graphics condition relative to the text, graphics & audio condition. Unequal sample sizes, which ranged from 19-34 participants (see Table 5.5.), could have had an impact on the significance of results. Therefore, any inferences regarding causality must be made with caution.

5.3.5 Group Comparison Summary

Data were analysed separately for three curriculum disciplines: science–topic heart; geography–topic map; and English–topic poem, to determine answers to the research question, “Does cognitive load affect learning outcome?” The additional factors of gender, WM, and motivation were explored to determine answers to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of cognitive load on learning outcome?” At times sample sizes were small and uneven and CAT scores not consistent across groups which could have influenced the outcome of results. Therefore, conclusions were drawn with caution. A summary of the levels of significance for each curriculum topic can be viewed in Table 5.20.

In answer to the research question, “Does cognitive load affect learning outcome?” findings confirmed, in the heart curriculum topic, that the retention of information for students was best in the delayed testing time, i.e., four to six weeks after being presented with the learning material, in the text, graphics and audio condition relative to the text-only condition. Results confirmed that a high cognitive load was better for students to study from in order to be able to recall information over an extended period of time than a low cognitive load. In the map curriculum topic, students recalled more information if the learning material was a combination of graphics and audio. Results confirmed that the most effective cognitive load for the retention of information from a map for students, was the one condition that did not include text. In the poem curriculum topic, results confirmed that the least effective combination for students was the highest cognitive load combination of text, graphics and audio, relative to text-only, and text and graphics. Findings identified that the most effective cognitive load for the retention of information for students varied between the different subject curriculum disciplines.

In answer to the research question, “Do individual differences (gender, WM, motivation) moderate the effect of cognitive load on learning outcome?” findings confirmed gender did not impact on recall in the heart and poem curriculum topics. In the map curriculum topic, girls retained the most information in the graphics and audio condition after a period of four to six weeks, relative to any other condition. Text appeared to reduce the amount of learning that took place. No significant differences between high and low WM groups were found in the the analysis investigating the effect of cognitive load, in any of the curriculum areas, which indicated that WM did not moderate the effect of cognitive load on recall. With regards to motivation, in the heart curriculum topic low motivation students performed significantly better in the text, graphics and audio condition than in any other condition–a high cognitive load appeared to engage students.

Table 5.20.

Summary of Significant Main Effects and Interactions for Each Curriculum Topic (Heart, Map, Poem) for Cognitive Load (Text-only; Text& Graphics; Graphics &Audio; Text, Graphics & Audio) with Gender, WM, and Motivation Presented Separately

Variables	Simple Main Effects & Interactions	Heart Task	Map Task	Poem Task
	Time	***	***	***
	Prior Knowledge	***	-	-
	Cognitive Load	-	**	**
	Time x Prior Knowledge	-	-	-
	Time x Cognitive Load	*	-	-
Gender	Time	***	-	***
	Prior Knowledge	**	-	-
	Cognitive Load	-	-	**
	Gender	-	-	-
	Time x Prior Knowledge	-	-	-
	Time x Cognitive Load	*	-	-
	Time x Gender	-	***	-
	Cognitive Load x Gender	-	*	-
	Time x Cognitive Load x Gender	-	-	-
WM	Time	***	***	**
	Prior Knowledge	***	-	-
	Cognitive Load	-	-	-
	WM	*	-	*
	Time x Prior Knowledge	-	-	*
	Time x Cognitive Load	-	-	-
	Time x WM	-	-	-
	Cognitive Load x WM	-	-	-
	Time x Cognitive Load x WM	-	-	-
Motivation	Time	***	-	-
	Prior Knowledge	**	-	-
	Cognitive Load	-	-	-
	Motivation	-	-	-
	Time x Prior Knowledge	-	-	-
	Time x Cognitive Load	-	-	-
	Time x Motivation	-	***	**
	Cognitive Load x Motivation	*	-	-
	Time x Cognitive Load x Motivation	-	*	*

*$p < .05$. **$p < .01$. ***$p < .001$.

In the map curriculum topic, within an hour after being presented with the learning material, it was found that high motivation students retained more information than low motivation students in the graphics and audio condition, and high motivation students performed better in the graphics and audio condition relative to the text, graphics and audio condition. After a period of four to six weeks after being presented with the learning material, it was found that high motivation students retained more information about the map in the graphics and audio condition relative to the text and graphics condition–audio and graphics appeared to be important components for retention in the map curriculum topic. In the poem curriculum topic, four to six weeks after being presented with the learning material, it was found that low motivation students retained less information in the text, graphics and audio condition than in any other condition, which indicated that a high cognitive load was not an effective learning method for low motivation students. High motivation students retained more information than low motivation participants in the text, graphics and audio condition.

Performance was relatively low in the map curriculum topic in both testing times, and participants did not appear to retain the information they learnt in all of the cognitive load conditions, which indicated that the information was initially not strongly encoded, and/or the curriculum topic was too difficult. However, there is further dialog on this issue in the Discussion.

Overall, results indicated no one cognitive load combination was suitable for all three curriculum topics for all students, and the type of curriculum topic should dictate the cognitive load condition most effective for learning. Findings could inform the methodology of delivery for students in different subject disciplines in the future.

5.4 Specific Characteristics of Participants (Age–Years 7, 8, and 9)

Participants represented a wide age range of students (11-14 years) within KS3, and age could have had an impact on recall. Analyses of variance were carried out to investigate whether any main effects or interactions existed between cognitive load, year group (7, 8, 9), and recall in both testing times (immediate, delayed) in each of the three topic curriculum areas (heart, map, poem). Results can be viewed in Table 5.21.

There were no Year 9 participants in the text and graphics condition for the heart curriculum topic, which was a result of assigning groups to different conditions in the mode of presentation (paper, M-technology) analysis to ensure equal sample grouping– and this may have had an impact on results in cognitive load the data analysis. The text and graphics sample group consisted of Year 7 and Year 8 students.

Table 5.21.

Analyses of Variance Examining the Difference Between Cognitive Load Scores (Text-only, Text & Graphics,Graphics & Audio, Text, Graphics & Audio) in Two Testing Times (Immediate, Delayed) in Each of the Three Curriculum Areas (Heart, Map, Poem), with Data for Years 7, 8, and 9 Presented Separately.

Task	Year	Cognitive Load	*N*	Time 1 *M (SD)*	*F*	*p*	Time 2 *M (SD)*	*F*	*p*
Heart	7	Text-only	15	8.73 (2.25)	1.87	.146	4.53 (3.18)	1.81	.156
		Text & Graphics	17	9.82 (2.58)			5.88 (2.85)		
		Graphics & Audio	14	9.71 (1.54)			5.07 (2.84)		
		Text, Graphics & Audio	13	8.15 (2.64)			3.54 (3.28)		
	8	Text-only	18	11.22 (2.73)	1.19	.319	5.17 (2.55)	2.35	.080
		Text & Graphics	18	10.17 (2.77)			5.00 (2.64)		
		Graphics & Audio	22	10.68 (2.06)			5.95 (2.82)		
		Text, Graphics & Audio	18	10.39 (1.72)			6.94 (2.01)		
	9	Text-only	35	10.83 (2.28)	1.28	.284	5.23 (3.07)	7.52	.001 **
		Text & Graphics		-			-		
		Graphics & Audio	13	10.23 (2.09)			5.62 (2.26)		
		Text, Graphics & Audio	20	11.45 (2.06)			7.75 (2.45)		
Map	7	Text-only	9	3.56 (2.56)	4.32	.009 **	1.00 (0.00)	7.86	.000 ***
		Text & Graphics	14	4.50 (2.57)			2.14 (1.10)		
		Graphics & Audio	17	6.47 (2.58)			4.41 (2.35)		
		Text, Graphics & Audio	17	3.71 (2.42)			2.59 (2.18)		
	8	Text-only	17	4.82 (1.88)	0.79	.501	1.76 (1.60)	1.91	.136
		Text & Graphics	18	5.17 (3.59)			2.39 (1.91)		
		Graphics & Audio	17	6.35 (3.10)			2.76 (2.11)		
		Text, Graphics & Audio	21	5.76 (3.55)			3.24 (2.07)		
	9	Text-only	32	4.94 (2.56)	1.16	.333	2.34 (1.43)	4.64	.005 **
		Text & Graphics	13	3.69 (1.75)			1.46 (0.88)		
		Graphics & Audio	13	5.15 (1.77)			3.15 (1.52)		
		Text, Graphics & Audio	10	4.50 (2.46)			1.50 (1.18)		
Poem	7	Text-only	18	8.94 (1.77)	1.40	.250	9.50 (1.72)	2.25	.091
		Text & Graphics	21	9.29 (1.55)			9.49 (2.31)		
		Graphics & Audio	16	8.25 (1.81)			7.94 (1.73)		
		Text, Graphics & Audio	15	9.27 (1.22)			8.73 (1.83)		
	8	Text-only	21	9.33 (1.43)	3.13	.031 *	7.95 (2.44)	3.25	.027 *
		Text & Graphics	15	9.00 (1.46)			9.47 (1.73)		
		Graphics & Audio	19	10.11 (1.24)			9.42 (1.54)		
		Text, Graphics & Audio	16	8.69 (1.62)			8.06 (1.98)		
	9	Text-only	19	9.79 (1.75)	9.78	.000 ***	8.84 (2.01)	2.83	.045 *
		Text & Graphics	20	10.25 (1.29)			8.65 (1.35)		
		Graphics & Audio	11	8.00 (1.67)			8.09 (1.92)		
		Text, Graphics & Audio	17	8.00 (1.66)			7.06 (2.61)		

Note. Year 7–mean age 12.00; Year 8–mean age 13.09; Year 9–mean age 14.07. T=Text-only, TG = Text & Graphics, GA = & Audio, TGA = Text, Graphics & Audio. Time 1 = immediate testing scores, Time 2 = delayed testing scores. **p* < .05. ***p* < .01. ****p* < .001.

ANCOVA analysis in the heart curriculum topic identified a significant main effect of Year, $F_{(2, 191)} = 8.58$, $MSE = 6.81$, $p < .001$. The interaction between time and cognitive load was not significant a, $F_{(3, 54)} = 8.58$, $MSE = 6.81$, $p < .001$, so data were collapsed and the average of the two testing times used. Year 9 students achieved significantly higher results, $M = 8.17$, $SE = .24$, than Year 7s, $M = 7.38$, $SE = .25$, and Year 8s, $M = 7.72$, $SE = .22$. There was also a significant interaction between Year and Cognitive Load, $F_{(5, 191)} = 3.69$, $MSE = 3.41$, $p < .01$. Bonferroni's adjusted means test investigated the pairwise differences between the cognitive load groups, while taking into account Type 1 errors. Pairwise comparisons identified that Year 9 students retained significantly more information in the text, graphics and audio condition, $M = 9.61$, $SE = .37$, than Year 7 students, $M = 6.27$, $SE = .48$. Year 8 students, $M = 8.29$, $SE = .41$, also retained significantly more information than Year 7 students, $M = 6.27$, $SE = .48$. No other pairwise comparisons were significant. Analysis demonstrated, in this instance, that there was a significant difference in recall scores between Year 9 students and Years 7 and 8 students. Both Year 8 and 9 students recalled more information in the highest cognitive load condition, i.e., text, graphics and audio, than Year 7 students.

In the map curriculum topic, ANOVA analysis identified there were no significant main effects or interactions that included Year groups. However, ANCOVA analysis in the poem curriculum topic identified a significant interaction between Year groups and cognitive load, $F_{(6, 195)} = 2.67$, $MSE = 1.76$, $p < .01$. Bonferroni's adjusted means test investigated the pairwise differences between the cognitive load groups, while taking into account Type 1 errors. Pairwise comparisons identified, in the graphics and audio condition, that Year 8 students retained significantly more information, $M = 9.77$, $SE = .35$, than Year 7 students, $M = 8.09$, $SE = .37$, and Year 9 students, $M = 8.04$, $SE = .45$. In the text, graphics and audio condition, Year 7 students retained significantly more information, $M = 9.08$, $SE = .40$, than Year 9 students, $M = 7.54$, $SE = .37$.

Findings identified significant differences in recall across year groups in the heart and poem curriculum topics. In the heart topic, Year 7 students recalled significantly less information than Year 8 and Year 9 students in the highest cognitive load condition (text, audio & graphics). In the poem topic, Year 8 students recalled significantly more information than Year 7 and Year 9 students in the graphics and audio condition, and Year 7 students retained more information than Year 9 students in the highest cognitive load condition (text, graphics & audio). Even though findings were not consistent across year groups, findings were not inconsistent across all year groups and across all curriculum topics, so, for the purpose of the current study, it must be noted that age may have been a potential influence of learning outcome.[2]

5.5 Chapter Summary

Statistical analyses were conducted to determine anwers to the research questions, "Does cognitive load affect learning outcome?" and "Do individual differences (gender, WM, motivation) moderate the effect of cognitive load on learning outcome?" At the start of the chapter, the general characteristics, i.e., descriptive statistics, of students were identified: The mean age and distribution of boys and girls was provided for each sample group (see Table 5.1., Table 5.2), and the means and standard deviations of student scores in each of the three curriculum topics (heart, map, poem) in immediate and delayed testing times (see Table 5.3., Table 5.4., Table 5.5.) were also presented. The distribution of students' cognitive ability (CAT) scores across cognitive load conditions was also provided. CAT scores identified the consistency of the cognitive capacity of students across experimental groups as inconsistent groups may have effected the findings.

In answer to the research question, "Does cognitive load affect learning outcome?" findings confirmed, in the heart curriculum topic, that a high cognitive load, i.e., using text, graphics and audio, was better for students to study from to be able to recall information over an extended period of time (four to six weeks), relative to text-only. In the map curriculum topic text appeared to moderate the affect of cognitive load in learning outcome, as the most effective condition for learning was the one condition that did not contain text, i.e., audio and graphics. In the poem curriculum topic, results confirmed that the least effective combination for students was the highest cognitive load condition, i.e., text, graphics and audio (relative to text-only, and text and graphics). Findings identified that the most effective cognitive load for the retention of information for students varied between the different subject curriculum disciplines.

In answer to the research question, "Do individual differences (gender, WM, motivation) moderate the effect of cognitive load on learning outcome?" findings confirmed gender did not have an impact on recall in the heart and poem curriculum

---

[2] As the research took place in an international school where 33% of participants' home language was not English, it was important to determine whether language had an influence on learning outcome. Analyses confirmed no significant main effects or interactions involving language (see Appendix R).

topics. In the map curriculum topic, girls retained the most information in the graphics and audio condition after an extended period of time, i.e., four to six weeks, relative to any other condition. No significant differences between high and low WM groups were found in the the analysis investigating the effect of cognitive load, in any of the curriculum areas, which indicated that WM did not moderate the effect of cognitive load on recall. In the heart curriculum topic low motivation students performed significantly better in the text, graphics and audio condition than in any other condition–a high cognitive load appeared to engage students.

Analysis of variance was later performed on data from students in Year 7, 8, and 9 to determine whether recall was consistent across year groups, i.e., whether age effected learning outcomes (see Table 5.21.). In certain instances, there were significant interactions between year groups and cognitive load, but as these significant interactions were not consistent across all year groups and topics, age may have had an influence on learning outcome.

Findings identified that when preparing lessons, educators should consider the curriculum topic and the type of information that needs to be learnt and plan their resources accordingly, e.g., science topics should use a combination of text and graphics, geography maps should include graphics and audio, and English poems should use text-only or a combination of text and graphics but not text, graphics and audio.

Chapter 6. Discussion and Conclusion

In this chapter, findings from the analyses of three separate curriculum topics relating to the three curriculum subjects: science–topic heart; geography–topic map; and English–topic poem, are discussed in relation to the research questions, “Does mode of presentation (paper, M-technology) affect learning outcome?”, “Does cognitive load affect learning outcome?” (text-only; text & graphics; graphics & audio; text, graphics & audio), and, “Do individual differences (gender, WM, motivation) moderate the effect of mode of presentation and cognitive load on learning outcome?” Conclusions are drawn from the findings of the current study, limitations are identified, and suggestions for future research acknowledged. A summary of the findings can be viewed in Table 4.18. and Table 5.20.

6.1 Do mode of presentation and individual differences affect learning outcome?

Mode of presentation compared the effectiveness of paper and M-technology resources on recall in three curriculum topics, in two testing times, to provide an answer to the research question, “Does mode of presentation (paper, M-technology) affect learning outcome?” A summary of levels of significance for mode of presentation can be viewed in Table 4.18. Analysis revealed no significant effects of mode of presentation in immediate or delayed recall in each of the three curriculum topics. With regards to individual differences, a difference was found in the heart curriculum topic where males performed better on paper-based than M-technology-based resources. No significant differences were found for WM or motivation between paper- and M-technology-based resources in any of the three curriculum topics.

According to Clariana and Wallace (2002), a test-mode effect should exist, and learning on either paper- or technology-based resources should lead to differences in performance. This hypothesis was not supported by findings in each of the three curriculum topics in the current study. Andrews and Haythornthwaite (2007) recognised potential issues with social implementation and differing teaching practice between paper- and technology-based learning. However, findings from the heart, map, and poem curriculum topics in the current study did not identify any social issues or different teaching practice, as there was no difference in outcome between paper or M-technology resources. Finding would indicate that the progression from paper- to technology-based learning was successful.

A level of significance was achieved with regards to time, during which more information was recalled in the first testing time than the second. This could be accounted for by Baddeley and Hitch's (1974) model of memory, whereby graphical information was managed by the visuo-spatial sketch pad, and text by the phonological loop. The limited storage and processing capacity of memory, combined with the complexity of information needing to be processed, could have exerted too great a demand on WM resources (Sweller, 2008), which could account for students recalling information in the short-term, but not over a longer period of time. Mayer's CTML assumption of active processing would explain the process of learning in the map curriculum topic as a blending of the information obtained through dual channel input, i.e., the spatial layout of the map and lexical street names were integrated with prior knowledge activated from LTM (Mayer, 2005; Mayer & Estrella, 2014; Paivio, 2007). Greater activation from LTM occurred in the short-term, therefore, more information was remembered in the earlier testing time.

The current study's findings can also be accounted for by the embedded-processes model of memory (Cowan, 1995, 1999), in which the amount of attention and focus paid to the content of the learning material resulted in a high level of LTM activation, which then transferred to the short-term recall of information when answering questions in the immediate testing time. The same amount of focus and attention may not have been present in the delayed recall testing time, i.e., not as much attention paid to the available cues, such as wording of the questions, therefore, the level of activation of LTM was not as great which resulted in less information being remembered.

The design of the resources could also have contributed to learning in the short-term. In both paper and M-technology resources, the content in the heart, map, and poem curriculum topics were formatted to aid recall. In the text-only condition, text was spaced consistently both horizontally and vertically, and key points were formatted to be bold. In the text and graphics condition, labels and text were situated close to the graphics they were describing (Abeysekera & Dawson, 2015; Norman, 2013). This organisation could have aided students to read faster, thereby supporting the retention of information (Hartley in Jonassen, 2004b).

Findings indicated that prior knowledge modulated the effects of mode of presentation in the heart curriculum topic, but not in the poem curriculum topic, even though students in the poem curriculum topic (N = 23) identified they were familiar with the story. Prior knowledge in the heart curriculum topic could have had an impact on learning outcome due to LTM becoming activated by word cues in the prior knowledge questionnaire (Mandler, 1980). In addition, significant results in the heart curriculum topic could be accounted for by the embedded-processes model (Cowan, 1999), in which prior lexical knowledge could also have helped to create a paired association between previously learnt information and new information, e.g., right ventricle and left ventricle ('right' and 'left' were the previously learnt information, and 'ventricle' the new information) semantically binding the two in the short-term (Savill et al., 2017). This process reduced the demands on WM so there was more capacity for processing the incoming information, which could explain why students recalled information immediately (Sweller, 2010). However, this paired association was of not long duration, the prior knowledge students had initially demonstrated was lost, along with the new information with which students were presented. Possibly, initial lexical familiarity was the cause of lack of consolidation–using familiar words may not have allowed permanent binding to occur, and so, after a short period of time, context and content were lost.

Students may have been unable to activate prior knowledge over a longer testing time due to decay, lack of rehearsal, and the influences of motivation and emotions, i.e., the mood-congruity effect (Baddeley & Hitch, 1974; Bower, 1981; Ebbinghaus, 1923/2013; Moreno, 2006). Language attrition may also have occurred, due to infrequently used words taking longer to recall,

Emerson and MacKay's (2011) question as to whether an Internet environment could result in students engaging less effectively with learning resources was not supported in the current study, as there were no significant differences in retention between paper or M-technology resources in each of the three curriculum topics. In addition, claims that technology enhances learning were not supported (Enriquez, 2010; Kalyaga & Liu, 2015; McMahon et al., 2016), as well as claims that technology can be detrimental to learning (Cromack, 2008; Emerson & MacKay, 2011; Wästlund et al., 2005). Findings from all three curriculum topics concurred with research that found no difference to learning outcome between paper and technology (Cakir & Simsek, 2010; Daniel & Woody, 2013; El Zein et al., 2016; Khoshsima et al., 2017; Mayer et al., 2007; Nicolli, 2015). Yet, Emerson and MacKay acknowledged the difficulty in identifying the best mode of delivery, as research has not taken into account other possible contributing factors, such as the reading age of students (Clark, 2013), or environmental conditions

(Plomin & Petrill, 1997). The current study did not take these factors into account, and reading age and environmental factors would need to be measured in future research.

Significant gender differences were found in the heart curriculum topic, in which boys showed higher retention using paper-based resources compared to M-technology resources. Findings from this curriculum topic did not align with research conducted by Clariana and Wallace (2002), who found no gender differences between paper- or computer-based conditions, and Papastergiou (2009), who found that outcome for boys and girls in secondary school computer science classes did not differ significantly. Findings also did not concur with Mangen et al.'s (2013) research in which students' comprehension was improved in a paper-based condition compared to technology-based condition. In the map and poem curriculum topics, there was no significant difference to learning outcome between boys and girls in each of the mode of presentation conditions (paper, M-technology) within an hour of being presented with the learning material as well as after four to six weeks.

With regards to WM, findings confirmed both modes of delivery were successful for all WM groups in both testing times for each of the three curriculum topics. This was in contrast to St Clair-Thompson and Gathercole (2006), who found WM influenced academic performance negatively, in that students who achieved low WM scores performed lower than expected in assessments. Findings also did not support Jabr's (2013) theory of intuitive navigation with location of text on a page that stated monitors interfere with the mapping process and affect WM function. Regarding Andrews and Haythornthwaite's (2007) question as to whether WM is enhanced through the use of technology or whether paper-based forms are best, results from the current study would answer, and say that the effect on WM would be the same for both paper and technology resources. In the poem curriculum topic, students with high WM performed better than the students with low WM, but not in relation to mode of presentation. Viewing one frame at a time on M-technology resources or multiple frames on one A4 sheet of paper, indicated that the design of the materials may have aided recall in the short-term for students with high WM, during which consistent structure and layout of the testing material reduced extraneous data processing and increased WM function (Lynch & Redpath, 2014; Moreno, 2007; Norman, 2013; Wong et al., 2012).

With regards to motivation, there was no significant difference to learning outcome between high motivation and low motivation students in both paper and M-technology conditions. In the heart curriculum topic students with high motivation outperformed students with low motivation, but this was not relative to mode of presentation. There must, therefore, be an alternative explanation for the achievement of highly motivated individuals. Within education, science is considered a core subject (Krapohl et al., 2014), and highly motivated students may have been more motivated to achieve in this topic on whichever medium the material was presented. Research has shown that newly implemented strategies have promoted science within education, thereby raising the status of the subject (Labrie et al., 2003). With the increased status of sciences, students may have perceived science to be more important than English, and were, therefore, more motivated in the heart curriculum topic than the poem curriculum topic. Also, the individual topics within the different subject disciplines may have been perceived differently by students. The heart topic is taught in biology in schools within regular lessons, and students may have believed this to be more important than what may have been perceived as a childish poem that would have no relevance to the student's future education.

Research has found that technology motivates and engages students more than paper, which has a positive effect on learning (Gabrielle, 2003; Mistler-Jackson & Songer, 2000). The current research did not support these findings. Intrinsic motivation may have prompted students in both modes of presentation, in each of the three curriculum topics, to have equal levels of engagement and motivation (Deci & Ryan, 1985). The current study does not support the earlier findings of Dede et al. (2005, April), who found using technology increased learning and engagement, and Driessen et al. (2007) who found Internet-based work increased students' motivation.

From a pedagogical perspective, an explanation for the lack of significance in performance between modes of presentation can be accounted for by the recent shift in methodology, whereby students have become independent learners with enhanced metacognition and problem-solving skills, focusing on high performance learning (Eyre, 2016; Herrington et al., 2008; Looi et al., 2010). This attitude to learning could reduce the effect the mode of presentation has on outcome, as students approach material presented in any form with the same mind-set, i.e., one of "can do."

In summation, findings indicated both paper and M-technology methods of delivery can hold a place in schools, and that the design of resources should be an important consideration. Findings also established that without revisiting information, there is a lack of consolidation and information can be lost. In addition, the shift in education methodology is beginning to reflect the practice of students out of school with regards to the use of smart technology (CISCO, 2017; Machin & Vignoles, 2005), whereby learning is now interactive and engaging (Buckner & Kim, 2013). It is necessary for continued progression to take place to keep education relevant and up-to-date. However, the decisions of educationalists regarding this progression should be supported by empirical research.

6.2 Do cognitive load and individual differences affect learning outcome?

Cognitive load in the current study explored the effect of different cognitive load conditions (text-only; text & graphics; graphics & audio; text, graphics & audio), on recall in three curriculum topics in two testing times, using the medium of M-technology (iPads®). A summary of levels of significance for cognitive load can be viewed in Table 5.20.

With regards to the most effective cognitive load for performance, findings varied between the subject curriculum tasks. In the heart curriculum topic, students in the highest cognitive load condition (text, graphics and audio) retained more information than those in the lowest cognitive load condition (text-only). Findings from the poem curriculum topic identified that students in the highest cognitive load condition (text, graphics and audio) retained less information than those in the text-only, and text and graphics conditions. Results from the heart curriculum topic refuted Gathercole et al.'s theory (2016) that stated the higher the demands placed on WM, the higher the cognitive load and the less likely it would be for information to be recalled. However, findings from the poem curriculum topic supported this theory.

Craik and Lockhart (1972) proposed that a semantic orientated task has better recall than a non-semantic task as it produces a trace available to a greater number of retrieval cues. In the current study, the heart, a familiar object, could be considered a semantic oriented task that could have aided recall. Even though resources included non-semantic technical terms, e.g., pericardium, students still managed to perform better in the higher cognitive load condition. The heart findings supported Muller et al. (2008) who found a higher cognitive load promoted effort and learning. A key difference between the heart and poem curriculum topic was the design of the material–the heart curriculum topic's graphics included a still graphic of the body torso with just the heart process animated (no extraneous data processing), and the poem curriculum topic's graphics were static, but there was more extraneous detail to process. This supports Kalyuga and Liu's (2015) split-attention theory where, in the heart curriculum topic, visual cues, i.e., the animated process, directed the students' attention to the salient points to be retained, and in the poem curriculum topic the students' attention was split between the intrinsic and extraneous cognitive load (Sweller, 2008).

In the map curriculum topic, students in the graphics and audio condition retained significantly more information than the text-only, and text and graphics conditions–a condition that did not contain any text. The map curriculum topic contained extended text that exceeded 400 words, and included non-semantic street names such as Haight Street. Findings supported Mayer's (2005) theory of active processing in CTML as students were able to blend spatial and verbal representations. Germane knowledge of map reading may have been activated from LTM and applied to the current scenario. Findings also supported Sweller's (2008) CLT that identified WM resources can be overloaded when extraneous activities are processed–in this case, processing text. However, even though there was a significant difference between cognitive load conditions, overall scores of the map curriculum topic were low in both testing times for all groups. This could be accounted for by the volume of data to be processed. According to Baddeley and Hitch (2000), capacity is limited by the number of items that can be articulated before their memory trace fades. The amount of data contained in the map curriculum topic, and lack of sufficient time to read through the material, may have hindered articulation and rehearsal, resulting in a weak memory trace and poor retention (Mayer & Estrella, 2014). Findings highlighted the importance of revisiting and reviewing information for it to remain in LTM, i.e., spaced repetition (Boettcher et al., 2018; Ebbinghaus, 1923/2013; Subrahmanyam, 2017).

Findings from the current study also support the embedded-processes model of memory (Cowan, 1995, 1999). In the heart, map, and poem curriculum topics, relevant prior knowledge from LTM was activated by cues provided in the testing material, and attention to these cues allowed learning to take place. The testing conditions that were shown to be statistically significant, may have had more semantic and non-semantic features activated from LTM, which created a stronger representation of the material. Students in the map curriculum topic, in which no prior knowledge existed except for frequently used words in the mental lexican, achieved lower marks than the other two curriculum topics. However, there may have been less representations for the map curriculum topic in LTM that could be called on. Individual differences, e.g., motivation, could have been influenced by the different levels of processing speed and experience of the students, and by the central executive paying more, or less, attention to specific data.

The individual difference of WM did not modulate the effects of cognitive load in all three curriculum topics. Findings in the heart curriculum topic confirmed students with high WM recalled significantly more information than students with low WM, but not relative to time or cognitive load. In the map and poem curriculum topics, analysis confirmed there was no significant difference between high WM students and low WM students in each of the cognitive load conditions in both testing times. Findings supported Sweller's (1988) premise that cognitive effort alone did not enhance learning outcome as consciously, or subconsciously, students must have paid attention to certain elements (road names, landmarks) which allowed them to manage cognitive resources. Findings did not support St Clair-Thompson and Gathercole (2006), who found students with low WM achieved lower in English assessments, as in the current research there were no significant differences between low and high WM students in the poem curriculum topic. The extended amount of script in the current research material concurred with Gorin and Embretson (2006) who found, in text-based research, that the greater the amount of text, the greater the demands on WM; both high and low WM students in the current study experienced cognitive load demands great enough to prevent learning over an extended period of time.

There were significant differences in motivation in all three subject curriculum areas. In the heart curriculum topic, low motivation students performed significantly better in the highest cognitive load condition (text, audio, and graphics) relative to the text-only condition. In the delayed recall testing time in the poem curriculum topic, low motivation students performed less well in the highest cognitive load condition relative to the text and graphics condition. In the poem curriculum topic the high motivation students performed better than the low motivation students in the highest cognitive load condition (text, graphics and audio). In the map curriculum topic low motivation students performed better in the graphics and audio condition relative to the text-only condition–low motivation students were more engaged when graphics were introduced. Findings could inform teaching methodology for low motivation students who are not engaged with learning. Findings support Moreno's (2006) Cognitive-Affective Theory of Learning with Media theory (CATLM) that acknowledged motivational factors affect learning. In the current study, it would appear that high and low motivation students increased or decreased their cognitive engagement through self-regulation. Mayer (2014) recognised motivation as a contributory factor for learning, but only if the design did not include distracting or extraneous elements; the heart curriculum topic did not have any superfluous information to process, but the poem and maps curriculum topics contained additional nonessential graphics. The current study recognised the importance of Spiegel and Rodríguez's (2016) criteria, that stated for M-technology to become a tool for learning, it is important to include the personal involvement of students, which includes motivation.

6.2.1 The science heart curriculum topic

In the science curriculum topic, results confirmed that using both visual and auditory channels in the text, graphics and audio condition, was better for long-term recall compared to text-only. Students were able to remember complex language and processes over time, when information was presented in a multimedia format. This was contrary to Mayer's (2009) CTML and Sweller et al.'s (1998) CLT, that proposed WM dual channel processes can be overloaded if presented with too much information–students were able to remember lengthy text, with a span greater than Miller's (1956) 7 (± 2), Cowan's (2001) 4 chunks, or Baddeley and Hitch's (2000) 16 related words in sentences. Results suggested that Miller, Cowan, and Baddeley and Hitch's memory chunk theories may only apply to short-term memory recall. According to Repovš and Baddeley (2006), visual input may be transferred from the visuo-spatial sketchpad to the phonological loop through recoding into a phonological form through articulatory rehearsal. As two forms of input were directed towards the visuo-spatial sketchpad (text, graphics), and one towards the phonological loop (audio), rehearsal of the text and graphics in the visual-spatial sketchpad, may have transferred some of the processing to the phonological loop, thereby utilising WM's capacity more efficiently.

Methodology and design may also have contributed to the success of the heart's high cognitive load condition. The dynamic graphics did not include extraneous, irrelevant activities to distract attention, and the pace of the animation was measured (Kalyuga & Liu, 2015). Students could also control the pace by forwarding, pausing, and rewinding. The split attention effect (Mayer & Moreno, 2003) had no impact, as the material was cohesive, with text appearing close to objects the same time as the audio played, which acted as attention cueing. The current study confirmed the findings of Yung and Paas' (2015a) research of cued animation of the circulatory system, in which aiding students to extricate relevant information resulted in better performance. The animation was limited to the processes heard in the audio, which supported Ayres and Paas' (2007) argument that animation can promote learning when designed to highlight relevant information.

As the results of the comparison between paper and M-technology were not significant in the short-term, a small amount of consolidation must have occurred over time, and students had the opportunity to reflect, either consciously or subconsciously on the heart material, thereby rehearsing it and strengthening the memory trace (Yeh, 2009). It has been identified that technology supports reflection when guiding questions are given in advance (Kori at al., 2014). The prior knowledge assessment questions, given before the testing condition, may have afforded this guidance. Therefore, the use of iPads®, together with priming questions, could have contributed to the consolidation of information.

In the delayed recall testing time, students who had demonstrated prior knowledge forgot some of their prior knowledge, even after being presented with cues in the form of questions. These results indicated that information in LTM may be lost even when the information is reinforced. An explanation could be that cue availability for the retrieval of information was influenced by the internal emotional state of students, and prevented the information from being recalled (Bower, 1981; Spear & Riccio, 1994). However, this is just speculation as the emotional state of students was not measured in the current study–further research would need to be conducted to confirm or refute this explanation.

Findings confirmed there was no significant difference in learning outcome between boys and girls in each of the cognitive load conditions in the immediate testing time and delayed testing time in the science curriculum topic. One could argue that Bevilacqua's (2016) explanation for gender differences, i.e., WM processes socially germane data (visuo-spatial for boys, language for girls), is becoming negated over time with the social pressure placed on girls to achieve in the sciences (Labrie et al., 2003). Social roles are changing, and girls are now encouraged to pursue sciences. Methodology is also changing, from teachers supporting boys more, to one of gender equality and practical application (Labrie et al., 2003; McGinnis & Tippins, 2001). Perhaps these changes are starting to take effect.

Motivation modulated the effects of cognitive load in the heart curriculum topic. Analysis confirmed low motivation students learnt better in the highest cognitive load condition (text, graphics and audio), in relation to the text-only condition. The higher cognitive load may have made the material more interesting for low motivation students, and, by engaging them more, established intrinsic motivation (Abeysekera & Dawson, 2015). The findings did not support Moreno's (2010) position in the cognitive-affect theory, that stated positive thoughts and belief would lead highly motivated students to achieve, or Sweller et al.'s (1998) CLT that did not recognise a connection between cognitive load and motivation. Lin et al. (2016) found reducing extraneous load through cueing in an interactive human cardiovascular research curriculum topic resulted in improved intrinsic motivation and learning–the current research confirmed these findings. It would seem the split-attention effect (Homer et al., 2008) was prevented through the design of the material, by taking into account the temporal and spatial contiguity principles (Mayer, 2009). Motivation did not modulate the effects of cognitive load for highly motivated students, and there was no significant differences in learning outcome between high motivation students and low motivation students in each of the cognitive load conditions. However, high motivation students performed better than low motivation students overall.

In summation, findings supported Meyer's (2005) multimedia principle, in which learning improved with combined words and pictures than from words alone, but not the redundancy principle in which learning should improve with reduced input. Combined text, graphics and audio was the best condition for learning over time, and consolidation may have occurred through rehearsal and reflection. Results indicated that the chunk effect was limited to short-term recall, and is negated over time with the correct methodology and combination of channels of input. With regards to individual differences, there were no gender differences, students with high WM recalled significantly more information in the heart curriculum topic than students with low WM, and students with low motivation learnt best in the highest cognitive load condition. Results revealed that the methodology for delivering information about the heart should include a high cognitive load, with three-dimensional animated graphics, delivered on M-technology and presented with initial guiding questions–with no extraneous information. This method of delivery should engage students with low motivation.

6.2.2 The geography map curriculum topic

In the geography curriculum topic, there was a significant difference in recall, in which students remembered the most information in the graphics and audio condition relative to the text-only, and text and graphics conditions. Time was not a significant factor. Results confirmed that the most effective cognitive load for the retention of spatial information, was the one condition that did not include text. These findings supported Mayer's (2009) CTML and Sweller et al.'s (1998) CLT. The most effective condition utilised two channels for processing information (Paivio, 2007), which may have allowed more information to be processed in WM as it was distributed between the phonological loop and visuo-spatial sketchpad, thereby not overloading one system. Results supported Sweller's (2008) modality effect for long-term memory encoding, that stated visual and auditory resources presented together (graphics and audio) are better than just a visual mode (text and graphics).

Findings from the current study concurred with Yung and Paas (2015b), who identified the importance of including graphics in resources, as they found combining graphical visual representations with text in a mathematics curriculum topic reduced extraneous cognitive load–using graphics and audio in the current study, may also have reduced extraneous load. Therefore, graphics can be viewed as one of the modes necessary to aid WM to build a representation of the elements to be processed. Yung and Paas identified single mode conditions were not successful, just as the text-only condition in the current study was not successful, confirming that dual mode processing (Paivio, 1990) is more effective than single mode processing. Within the current study, the static map graphic provided a constant element from which students could visually navigate while listening to the audio (Wong et al., 2012).

With regards to individual differences, differences in performance were identified within gender; girls performed significantly better in the graphics and audio condition in the delayed recall testing time, relative to the text-only and text and graphics conditions–it was a condition that did not include text, which reflects the overall outcome of the map curriculum topic. No significant gender differences were found between boys and girls. Bevilacqua (2016) claimed that males are strong in visuo-spatial processing and females in language processing. One could argue that within the audio-visual condition, the processing would differ for each gender, at which time males use the visuo-spatial sketchpad system and females the accompanying sound through the phonological loop system (Baddeley & Hitch, 1974). The end result was the same, as there was no significant difference in learning outcome between boys and girls in the graphics and audio condition. However, the information may have been processed differently. Further psychological research would need to be conducted, isolating each factor to determine whether this was indeed the case.

The interaction between motivation, cognitive load, and testing time was significant. Analysis revealed, in immediate recall, that students with high motivation performed better in the graphics and audio condition than students with low motivation. Also, in immediate recall, students with high motivation performed better in the graphics and audio condition relative to the text, graphics and audio condition. In delayed recall, students with low motivation performed better in the graphics and audio condition relative to the text–only condition, and students with high motivation performed better in the graphics and audio condition relative to the text and graphics condition. Findings supported Moreno's (2010) cognitive-affect theory, that states a positive attitude ensures cognitive resources are assigned to the curriculum topic and learning will take place. Findings, however, did not concur with Lin et al. (2016), who found learning can take place even if motivation is not present.

Sweller et al.'s (1998) CLT identified that mental effort is necessary for learning to occur, and recognised the importance of a methodology that decreased extraneous cognitive load. A successful methodology for decreasing cognitive load in the current study was combining static graphics and audio with a human voice recording. Even though Moreno (2010) criticised the CLT for not taking into account individual differences, the methodology in the current study took into account Sweller's theory, and identified that the CLT was suitable for the individual differences of gender and motivation.

In summation, the best condition for short- and long-term spatial manipulation of a map was a combination of audio and graphics. Findings from the current study have supported Meyer's (2005) redundancy principle that states learning is improved with reduced input, and the multimedia principle that states learning is improved with combined words (audio) and pictures. These results inform the methodology of delivery to students with regards to map navigation curriculum topics.

6.2.3 The English poem curriculum topic

In the English curriculum topic, results confirmed that the highest cognitive load, i.e., text, graphics and audio, was less successful for learning relative to text and graphics. The findings supported Mayer's (2005) CTML and Sweller's (2008) CLT that identified WM resources can be overloaded when extraneous information is processed. CLT's split attention effect may have had an effect and prevented a strong memory trace being created, which resulted in the loss of information (Kalyuga & Liu, 2015). Even though the text contained relatively simple language, the sheer volume (516 words) may have overload WM, which may have prevented strong encoding (Baddeley & Hitch, 2000). Acoustically similar rhyming words may also have had made an impact on learning, with students focusing their attention on the rhyming aspect rather than the content, resulting in the loss of information (Logie, 1995).

No difference was found between boys and girls in both testing times. This could have been the result of boys processing information via the visuo-spatial sketchpad system and girls processing the accompanying sound through the phonological loop system equally well (Baddeley & Hitch, 1974). Results did not support the PISA meta-analysis in which girls scored higher in reading (Stoet & Geary, 2013). If this was indeed the case, then girls would have outperformed boys significantly in at least one of the conditions that contained text in the English poem curriculum topic.

In delayed recall, analysis revealed that low motivation students retained less information in the high cognitive load condition (text, graphics and audio) relative to text and graphics, and that the high motivation students performed significantly better than the low motivation students in the text, graphics and audio condition. Findings did not support Hawlitschek and Joeckel (2017), who found that motivation did not influence learning. Findings from the current study's poem curriculum topic concurred with Homer et al. (2008), who found that greater motivation occurred with resources that had no animation, and which were narrated with a human voice.

In summation, even though dual processing may have taken place, cognitive overload may have prevented strong encoding, which resulted in reduced retention. There were no differences in learning between cognitive load conditions for gender or WM, however, cognitive load influenced the impact of motivation. Findings supported Meyer's (2005) redundancy principle, that stated learning should improve with reduced input, but not the multimedia principle, that stated learning should improve with combined words and pictures rather than from words alone. These results inform the methodology of delivery to students with regards to largely text-based curriculum topics, and identify the dangers to learning of cognitive overload.

6.2.4 Cognitive load summary

Examination of the findings across all three curriculum topics challenged Gardner's (1983) theory of multiple intelligences, in which learning styles are different for each child. Findings supported the dual channel processing theory of Paivio (2007) and the transferral theory of Repovš and Baddeley's (2006), in which data from the visuo-spatial sketchpad is transferred to the phonological loop through recoding into a phonological form through articulatory rehearsal. Results in the heart and poem curriculum topics confirmed extended data can be recalled, providing the material is presented to students in a format that leads to better encoding and retention. Therefore, design was identified as an important component of methodology, and keeping information germane through cues, students control, and teacher's instructions prior to learning, was essential for long-term retention. Graphics was also a key element to include in all types of curriculum topics. Findings from the current study also demonstrated that when the methodology is correct, gender differences are reduced and the equality of learning is improved. Motivation had an impact in all the subject curriculum curriculum topics, confirming material needs to be designed with high and low motivation students in mind. The current research confirmed that the nature of the curriculum topic should dictate the method of delivery, and that students who are not motivated may become engaged depending on cognitive load and design.

6.3 Limitations and Future Research

One insight arising from the research process has been that a dichotomous distinction between paper and M-technology as the manipulation for mode of presentation was perhaps too simplistic. A more fruitful approach might have been to include other variables, such as physical and mental interactivity as identified by Sweller (2010). The identification of low and high interactivity of elements within the testing material could have been used in both paper and M-technology resources. For example, in the current study there was little physical interactivity with the paper resources, but there may well have been a cognitive interactivity, i.e., immersion with the content of the data. In the M-technology condition, there was more physical interactivity as students could manipulate the start, speed, and pause of the material. However, students may not have been immersed in the content. Therefore, future research could focus on identifying different levels of interactivity in both paper and m-technology modes of presentation, to shed further light on the role of mode of presentation and learning.

According to the diffusion model (Giguère & Love, 2013), when a person tries to remember an event or recall knowledge, information may not be available immediately, i.e., the correct information may have to be retrieved from a myriad of information. This process takes time and if not given enough time to locate the right information from memory, a person may not be able to answer a question, which they may have done if given more time. According to the diffusion model, “it is impossible to make optimal decisions [recall memory] within finite time” (Giguère & Love, 2013, p. 7614). In the current study, the time allocated for assessing the recall of students (immediate testing time, delayed testing time) may not have been enough–students, who may have known the answers, may not have answered questions as they ran out of time. A solution would have been to extend the time given to students in the immediate and delayed testing times, to ensure they had enough time to recall the information from LTM.

The current study was a quasi-experiment where students were selected from an opportunity sample. Students, therefore, did not represent a truly randomised sample of the target population, but were part of a non-true randomised design. As a consequence, it is difficult to generalise the current findings to other situations. In addition, as a quasi-experiment, not all the variables could be determined in advance of the research taking place. Extraneous variables, such as temperature within rooms, external noise, etc., could not be controlled by the researcher, and these variable may well have impacted on the findings. An experiment is designed to identify cause and effect, and with all the factors identified with regards to the current study, causality inferences can not be made. One can possibly look at similar situations, i.e., international schools with a British curriculum, and see whether the current study provides some useful way to help other teachers decide how to create and structure their learning resources.

It is possible that students’ performance in the study may have been influenced by the Hawthorne effect, that is, where students are aware of the fact that they are participating in a study and attempt to anticipate expected responses, which might influence their performance. In the current study students may have expected that learning using the M-technology resource (iPad) may have been more successful than paper, and those on paper-based resources may not have tried as hard to recall information, or deliberately (or unconsciously) did not write down answer they knew. It is difficult to find a solution to such demand characteristics – one way would be to mis-inform students and tell them that the research was focusing on a different learning outcome. There are ethical issues of 'informed consent' with this approach, but as long as there is no physical or mental harm, it is an option to consider. To minimise the effects of such demand characteristics, the current research encouraged students to try their best at all times.

The research conducted within the current study examined learning within a school environment. It did not take into account additional factors, such as the home environment or social aspects that could support alternate theories such as emotions affecting learning (Mayer & Estrella, 2014). Additional factors could also account for why a student with a high WM ability underachieves, or a student with a low WM ability overachieves. This could be an area for future research.

Sustaining the complex relationship between the researcher, school, and parents, as well as collecting, marking, and analysing data, was a tremendous amount of work, and researcher fatigue was a reality (Clark, 2008). The researcher dealt with fatigue through efficient time management and support from the school. Participant fatigue could also have occurred as multiple data collection sessions took place; students were encouraged at the start of every session.

The sampling method of selecting students from one school, i.e., an opportunity sample, could be potentially biased. However, the school has an enrolment of students from over fifty nations, and that established a naturally occurring randomised sample, thereby demonstrating greater generalisability (Wang et al., 2009).

A theoretical link was made between WM and intelligence (Cowan, 2014; Halpern, 2013). Future research should include both intelligence and WM assessments for all students. Sample sizes would then be larger, and the results more reliable. In addition, results for mode of presentation and gender differences indicated that boys performed better on paper than technology. However, levels of intelligence were not determined in advance, so more intelligent boys could have randomly been placed in the paper condition. Future research should allocate students to groups through a matched pairs design.

The current study used a repeated measures method, as it researched recall and perception over two points in time with the same students. Perceptions and schemas change as individuals grow and gain experience (Abeysekera & Dawson, 2015; Veenman et al., 2006). Including a third time point may have provided the opportunity to compare performance between time points one and two, one and three, and two and three, to determine exactly how much recall varied. However, the research set out to determine recall with students within a specific age range, and extending the testing time may have placed some students outside the specified age. In addition, other factors would need to be considered, such as students leaving the school, content being revisited in the academic curriculum, and added students fatigue.

In general, students were keen participators. Yet, they were aware they were being observed, and may have suffered from performance anxiety. This may have resulted in a change of behaviour from the norm, which, in turn, could have influenced the outcome (Hanson, 1967). However, multiple data were collected from students over an extended time period, during which they may have gradually felt more at ease. Also, the natural school environment, in which students participated in normal school classes before and after the data collection, may have served to counterbalance any anxiety effects.

The reading speed of students could have had an impact on retention, as opposed to cognitive load. Slow readers would have taken a longer time to read through the material, perhaps not even reading through the whole text before the end of the timed condition. In future research, reading speeds would need to be measured, and additional time provided for slower readers.

Students may have developed an interest in the research topics, and revisited the material between testing times, or even read more around the subject (Cohen et al., 2017). To determine whether this was the case, in the final questionnaire students were asked whether they had reviewed the material between testing times. In the different curriculum topics (heart N = 19, poem N = 21, map N = 13), students admitted to accessing the material, from one to four times straight after the initial data collection to the day of the final assessment. Considering the large sample size for each curriculum topic (heart N = 311, poem N = 289, map N = 304), these numbers should not have been large enough to affect the overall results (Field, 2014).

With regards to generalisation, the data were collected from a multinational high-achieving international school, in which many students were adept in the use of M-technology (Mayer, 2009). However, conditions in schools may differ, e.g., teachers may hold on to traditional forms of delivery (Tondeur et al., 2017), and the current findings may not be directly applicable to those schools. However, comparisons with other similar schools could either support or refute M-technology as a methodology for improving learning, from which progressing schools could learn (Cohen et al., 2017).

Reflection, reviewing information, and spaced learning were discussed as possible explanations for consolidation (Ebbinghaus, 1923/2013; Kori et al., 2014; Procee, 2006; Subrahmanyam, 2017). However, additional theories, such as consolidation during sleep, and empirical evidence from neurological research, should be included in memory research. In addition, discourse on theories of forgetting, such as cognitive overload, the lack of rehearsal, motivation, and the emotional state of students, were provided (Bower, 1981; Deci & Ryan, 2000; Lowe, 2003; Mayer, 2009), but further factors, such as 'modern-day' interferences, e.g., using multiple devices, could be explored more in the future.

Results for the poem mode of presentation showed very little difference between scores gained in the immediate recall and delayed recall testing times. This could have been a design fault, and a ceiling effect could have potentially impacted on the results. Future research should determine the complexity of the testing material to prevent a ceiling effect from occurring.

A further limitation of the study was not conducting research that included qualitative data collection. Crookes (2013) acknowledged that social science research has been recognised as being inter-disciplinary, requiring different methods of research that encompass both objectivity and interpretivism, i.e., a "plurality of research methods" (p. 3) that discover how knowledge is attained. Wertz et al. (2011) acknowledged how well qualitative data analysis complements quantitative methods. Multiple methods allow complex educational research issues to be examined and encourages exploration of the problem from different perspectives, thereby gaining a better understanding of the research problem and avoiding bias of single data collection through internal cross-checking (Creswell & Creswell, 2017; Du, 2012; Gill & Johnson, 2010). Willig and Stainton-Rogers (2017) acknowledged the growing importance of qualitative data in informing researchers' conclusions. Quantitative methods allow significant interactions and effects to be determined (Johnson et al., 2007), and the qualitative data allows responses from students to enrich the data gained (Gill & Johnson, 2010). Qualitative data collection has also been identified as often missing from technology research in schools (Pérez-Sanagustín et al., 2017). Future research incorporating both qualitative and quatitive aspects has the potential to build on findings of the current research.

6.4 Implications for Practice

6.4.1 Mode of Presentation

In the current study there were very few significant findings with regards to mode of presentation and learning outcome. There was no difference to students' learning if they used a paper-based resource or a mobile technology device (iPad®) in the heart, map, and poem topics (see Table 4.18). The results were interesting and were not expected. With the ubiquitous nature of technology today (CISCO, 2017), one would have thought that technology would have had more of a significant impact on learning. Clariana and Wallace also (2002) acknowledged that a test mode effect should exist where paper and technology would not achieve the same results. However, the international school had embraced the use of technology for a few years before the research took place, and students used their own iPad® on a daily basis. Students were competent in using both textbooks and technology and the findings of the current study may have reflected this practice. It may not have mattered what subject curriculum topics were selected for the research, as it was likely that the findings would have been consistent across all subject disciplines.

A couple of points can be gained from the results. What the school can learn from this research is that it is important to have a balanced approach to the use of resources. Paper-based medium may be around for a while in some industries (law, schools–not all schools have access to technology), while other other industries are leaders of innovative technology (engineering, computer science). The school is preparing students to work in industries that utilise both paper-based and technology media, and should therefore continue to use both textbook and iPads® as it can be viewed as beneficial to medium-term gains for students. This is one perspective. Another perspective recognises, that as technology is not detrimental to learning, and with the ubiquitous nature of technology today, it provides an opportunity for teachers to extend students learning (and for students themselves), to progress and develop even more innovative technology-based teaching and learning resources.

There were significant results with regards to gender differences. Boys learnt better using paper-based resources than M-technology in the heart task, and girls retained more information than boys in the map task within an hour of the testing period. Findings have identified that gender needs to be considered with regard to learning resources. Teachers need to plan lessons that take individual students' needs into account, i.e., differentiated lessons in a science topic, where boys are given resources on paper and girls could be given the option on which media they would prefer to work. This would prepare students for the best possible learning outcome.

6.4.2 Cognitive Load

There were significant differences between cognitive load conditions in each of the three topics. In the heart topic students retained more information in the text, graphics and audio condition than the text-only condition. Findings identified in a science-based topic that students learnt better in a dual-coding higher cognitive load condition than a single chanel text-only condition (Paivio, 1990). Low motivation students also performed better in this condition than in the text-only condition. It would appear that a high cognitive load engaged low motivation students. This is a powerful message for schools as it identifies ways to help low motivation students to achieve.

In the map topic, students in the text, graphics and audio condition performed better than in the text-only and text and graphics conditions-this was consistent with girls too. High motivation students performed better in the graphics and audio condition compared to the text, graphics and audio condition within an hour of learning, and in the graphics and audio condtion relative to text and graphics after four to six weeks. Low motivation students also retained more information in the graphics and audio condition after four to six weeks, but relative to the text-only condition.

In the poem topic students in the text-only condition performed better than students in the text, graphics and audio condition. A high cognitive load appeared to be detrimental to learning in an English text-based topic. High motivation students performed better than low motivation students in the text, graphics and audio condition, and low motivation students achieved better in the text and graphics condition than in the text, graphics and audio condition.

Findings from the current study have identified that teachers can make an informed decision as to the best way to present information to students. The curriculum topic should determine the material students require to enhance their learning. Students, as stakeholders, can be made aware of the different cognitive load combinations and the effect they have on learning, so they can choose effective learning resources. However, subjects do not need to be limited to one cognitive load combination, e.g., geography does not only teach about maps, science does not only use labelled diagrams. The type of topic should determine the type of cognitive load, and any one department can use a variety depending on the information that needs to be learnt.

Important discoveries regarding low motivation students and strategies with regards to cognitive load and how they can become more engaged with learning has been identified. This is extremely useful information for schools to implement.

6.4.3 Summary

It has taken over 40 years to achieve Dewey (1929) and Skinner's (1965) vision, ensuring that learning is individualised, dynamic, and uses an intelligent machine that can assess student progress. The advent of technology has resulted in a transformation in educational methodology, moving from paper- to computer-based learning (Fojtik, 2015; Hartnell-Young & Heym, 2008). Initially this was not based on empirical evidence, and little research exists today comparing these two media in an educational setting. However, of the research that does exist, including the current study, there is more support for both paper- and technology-based methods of delivery for learning (Clariana &Wallace, 2002; Khoshsima et al., 2017). Yet, this is almost superfluous, as the ubiquitous nature of technology demands that it be incorporated into teaching practice and it would be disadvantageous to the future contribution of students to society if they were not supported on this medium (Cavus & Uzunboylu, 2009; Nedungadi & Raman, 2012; Wingkvist & Ericsson, 2013). In addition, the changing landscape of learning, which has become far more student led, can provide a greater variety of innovative ways for students to learn.

What has been identified in the current study is that learning styles are not necessarily fixed as determined by Riding and Douglas (2018) and Gardener's (1983) multiple intelligences. What has been determined is that learning depends on the type of content, i.e., text-based; linking names to objects and identifying systems and processes; or manipulating a map spatially, and that resources should be created and used to promote learning within that type of context–the cognitive load learning resource should suit the content. Future research should include additional types of curriculum topics, such as programming, to ascertain the correct cognitive load for discrete curriculum topics; this would identify additional topics and best ways to present resources which would result in an informed and customised teaching practice for individual subjects within schools.

6.5 Conclusion

The current study highlighted the relevance of a systems approach that takes into account human learning through information processing, as well as the tools with which to do it (Jonassen, 2004a). The critical realism approach of discovering existing structures and processes through empirical research sat well with activity theory that identified the external influences which support learning through daily action (Kaptelonin & Nardi, 2006; Morgan & Smircich, 1980; Sayer, 1992).

Moreno's (2006) CATLM acknowledged the importance motivation has in controlling WM processes during learning, and findings from the current study confirmed motivation played a significant role in recall in each curriculum topic. In addition to motivation identified by Moreno (2006), Mayer (2014) reinforced the significance of design via the redundancy principle. Design was also identified in the current study as an important consideration for managing cognitive load. Reducing extraneous material, and including visual signals and elements that learners can control (Abeysekera & Dawson, 2015; Norman, 2013), enabled the retention of germane information (Gillmor et al., 2015; Lynch & Redpath, 2014; Sweller, 2010). It has been acknowledged that giving students instructions and direction prior to the start of independent learning tasks leads to better recall (El Zein et al., 2016; Leinonen et al., 2016). Therefore, the variables of motivation, design, and verbal instructions, were recognised as important for the retention of information, and all three should be documented as interconnected influences that contribute to learning.

Models of learning recognise that WM has a limited capacity, and that cognitive load affects how much learning can take place (Cowan, 1999; Mayer, 2005; Moreno, 2006, Sweller, 2008). The current study recognised the importance of examining memory and learning alongside Mayer's multimedia and redundancy principles. The current study also confirmed that learning can take place with a high cognitive load, but the materials should include cueing, and reducing superfluous elements (Mayer, 2014; Sweller, 2010, Wong et al., 2012). Students were able to recall extended information, i.e., more than 16 word sentences (Baddeley & Hitch, 2000), in an educational setting; it was a true representation of situated learning, i.e., the classroom and content of the resources in the study reflected a realistic learning environment. Future research should explore further extended texts and graphics to determine the length of information that can be recalled successfully.

The current study set out to determine the impact of M-technology and cognitive load on learning in the light of contemporary resources and approaches adopted by an international school, in which M-technology had been adopted without any empirical evidence that this medium was effective in enabling students to retain information. Findings in the current study have identified that M-technology does not have a significant positive impact on the retention of information; however, it also does not have a negative impact. Therefore, using M-technology in an educational setting does no harm to learning. With regards to cognitive load, the current study recognised that the same cognitive load cannot be applied to all learning conditions. Results indicated that retention was improved by different combinations of text, graphics and audio, and that the characteristics of the task should determine the cognitive load that should be used.

Overall retention in the map curriculum topic was low. This low level of retention raises questions with regards to data consolidation. Students were given five minutes to read through the resources, and then asked to recall the information forty minutes later, and again four to six weeks later. A stand-alone five minute session may not have been enough time for students to create a strong memory trace with extended text and graphics, regardless of how much rehearsal occurred within that time frame. The low level of retention identified the importance of consolidation, and educationalists should recognise that students need multiple engagement to ensure that retention takes place. Teachers should be aware they need to revisit and revise every aspect of a topic, i.e., spaced learning, and not just present it once.

In conclusion, the current study warns educationalists to be careful about making claims for innovation without any data to support gains. The introduction of technology by educational institutions, should be supported by controlled studies to ensure that at the very best they do no harm, and teachers should be apprised of up-to-date research teaching practice to ensure they are supporting students to achieve their very best.

References

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. *Higher Education Research & Development*, *34*(1), 1-14. Retrieved from https://doi.org/10.1080/07294360.2014.934336

Ally, M., & Prieto-Blázquez, J. (2014). What is the future of mobile learning in education? *International Journal of Educational Technology in Higher Education*, *11*(1), 142-151. Retrieved from http://dx.doi.org/10.7238/rusc.v11i1.2033

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001) Development of Executive Functions Through Late Childhood and Adolescence in an Australian Sample, *Developmental Neuropsychology, 20*(1), 385-406. Retrieved from https://doi.org/10.1207/S15326942DN2001_5

Andrews, R., & Haythornthwaite, C. (Eds.). (2007). *Handbook of e-Learning Research*. London, UK: Sage.

Arthur, P. L. (2018). Engaging Collections and Communities: Technology and Interactivity in Museums. Retrieved from http://ceur-ws.org/Vol-2084/shortplus11.pdf

Ary, D., Jacobs, L. C., Irvine, C. K. S., & Walker, D. (2018). *Introduction to research in education*. Cengage Learning.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of learning and motivation*, *2*, 89-195. Retrieved from https://doi.org/10.1016/S0079-7421(08)60422-3

Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology*, *21*(6), 695-700. doi:10.1002/acp.1343

Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417-423. Retrieved from https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology*: *General*, *129*(1), 126.

Baddeley, A.D., & Hitch, G.J. (1974). Working Memory. *Psychology of Learning and Motivation, 8,* 47-89. Retrieved from https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A.D., & Hitch, G.J. (2000). Development of Working Memory: Should the Pascual-Leone and the Baddeley and Hitch Models Be Merged? *Journal of Experimental Child Psychology*, *77*(2), 128-137. doi:10.1006/jecp.2000.2592

Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement, 60*(3), 361-370. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/00131640021970592

Baylor, A., Ryu, J. & Shen, E. (2003, June). The Effects of Pedagogical Agent Voice and Animation on Learning, Motivation and Perceived Persona. In D. Lassner, & C. McNaught (Eds.), *Proceedings of ED-MEDIA 2003 - World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 452-458). Honolulu, Hawaii, USA: Association for the Advancement of Computing in Education (AACE). Retrieved from https://www.learntechlib.org/primary/p/13800/

Belmore, S. M. (1985). Reading computer-presented text. *Bulletin of the Psychonomic Society, 23*(1), 12-14. doi:10.3758/BF03329765.pdf

Bennett, S. (2012). Digital natives. In Z. Yan (Ed.), *Encyclopedia of Cyber Behavior: Volume 1* (pp. 212-219). United States: IGI Global

Borghans, L., Golsteyn, B. H., Heckman, J. J., & Humphries, J. E. (2016). What grades and achievement tests measure. *Proceedings of the National Academy of Sciences, 113*(47), 13354-13359.

Brinson, J. R. (2015). Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: A review of the empirical research. *Computers & Education, 87*, 218-237.

Bryman, A. (2016). *Social research methods*. Oxford university press.

Burgin, L., Ekström, M., & Dessai, S. (2017). Combining dispersion modelling with synoptic patterns to understand the wind-borne transport into the UK of the bluetongue disease vector. *International Journal of Biometeorology, 61*(7), 1233–1245. Retrieved from https://link.springer.com/content/pdf/10.1007%2Fs00484-016-1301-1.pdf

Bergson, H., Paul, N. M., & Palmer, W. S. (2004). *Matter and Memory*. New York: Dover Publications.

Bevilacqua, A. (2016). Commentary: Should Gender Differences be Included in the Evolutionary Upgrade to Cognitive Load Theory? *Educational Psychology Review,* 1-6. doi:10.1007/s10648-016-9362-6

Bhaskar, R. (2013). *A Realist Theory of Science*. Abingdon, Oxon: Routledge.

Biesta, G.J.J. (2017). *The Rediscovery of Teaching*. New York, NY: Routledge.

Billings, E. S., & Mathison, C. (2012). I get to use an iPod in school? Using technology-based advance organizers to support the academic success of English learners. *Journal of Science Education and Technology*, *21*(4), 494-503. doi:10.1007/s10956-011-9341-0

Boettcher, M., Boettcher, J., Mietzsch, S., Krebs, T., Bergholz, R., & Reinshagen, K. (2018). The spaced learning concept significantly improves training for laparoscopic suturing: a pilot randomized controlled study, *Surgical Endoscopy*, *32*(1), 154-159. Retrieved from https://doi.org/10.1007/s00464-017-5650-6

Bortolussi, M., & Dixon, P. (2013). Minding the text: Memory for literary narrative. In L. Bernaerts, L. Herman, B. Vervaeck, & D. de Geest (Eds.) *Stories and minds: Cognitive approaches to literary narrative* (pp. 23-37). Lincoln: University of Nebraska Press.

Boud, D., & Walker, D. (1998). Promoting Reflection in Professional Courses: The Challenge of Context, *Studies in Higher Education*, *23*(2), 191-206. doi:10.1080/03075079812331380384

Bower, G. H. (1981). Mood and Memory. *American Psychologist, 36*(2), 129-148. doi:0003-066X/81/3602-0129S00.75

Brady, T.F., Störmer, V.S., & Alvarez, G.A., (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences of the United States of America, 113*(27), 7459-7464. Retrieved from https://doi.org/10.1073/pnas.1520027113

Bruno, A. & Dell'Aversana, G. (2018) Reflective practicum in higher education: the influence of the learning environment on the quality of learning, *Assessment & Evaluation in Higher Education*, *43*(3), 345-358. Retrieved from DOI: 10.1080/02602938.2017.1344823

Bryman, A. (2013) *Quality Issues in Mixed Methods Research*. Presented at The White Rose Social Science DTC Second Annual Spring Conference, 7th May 2013, University of Leeds, UK.

Buckner, E., & Kim, P. (2013). Integrating technology and pedagogy for inquiry-based learning: The Stanford Mobile Inquiry-based Learning Environment (SMILE). *Prospects,* (44), 99-118. Retrieved from https://doi.org/10.1007/s11125-013-9269-7

Bunch, R. L., & Lloyd, R. E. (2006). The cognitive load of geographic information. *The Professional Geographer*, *58*(2), 209-220. Retrieved from https://doi.org/10.1111/j.1467-9272.2006.00527.x

Busher, H., & James, N. (2012). Ethics of research in Education (3rd ed.). In A. Briggs, M. Coleman, & M. Morrison (Eds.), *Research Methods in Educational Leadership and Management* (pp. 90-104). London: SAGE Publications Ltd. Retrieved from https://trove.nla.gov.au/version/222666830

Cakir, O., & Simsek, N. (2010). A comparative analysis of the effects of computer and paper-based personalization on student achievement. *Computers & Education, 55*(4), 1524-1531. doi:10.1016/j.compedu.2010.06.018

Calvo, R. A., D'Mello, S., Gratch, J., & Kappas, A. (Eds.), (2014). *The Oxford Handbook of Affective Computing*. USA: Oxford University Press.

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, *54*, 297–312. doi:10.1037/h0040950

Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research.* Ravenio Books..

Cardwell, M. (2014). *Dictionary of Psychology*. Abingdon, Oxon: Routledge.

Cassidy, T., & Lynn, R. (1991). Achievement motivation, educational attainment, cycles of disadvantage and social competence: Some longitudinal data. *British Journal of Educational Psychology*, *61*(1), 1-12. Retrieved from https://doi.org/10.1111/j.2044-8279.1991.tb00956.x

Cavus, N., & Uzunboylu, H. (2009). Improving Critical Thinking Skills in Mobile Learning. *Procedia Social and Behavioural Sciences*, *1*, 434-438. Retrieved from https://doi. org/10.1016/j.sbspro.2009.01.078

Chang, C. C., Liang, C., Chou, P. N., & Lin, G. Y. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? Perspective from multimedia and media richness. *Computers in Human Behavior, 71*, 218-227. Retrieved from https://doi.org/10.1016/j.chb.2017.01.031

Chen, S. C., Yang, S. J., & Hsiao, C. C. (2016). Exploring student perceptions, learning outcome and gender differences in a flipped mathematics course. *British Journal of Educational Technology, 47*(6), 1096-1112.

Chen, X., Jiang, M., Cai, L., Liang, L. L., Du, J., & Zhou, Y. (2017). Alignment Between the National Science Curriculum Standards and Standardized Exams at Secondary School Gateways. In L. Liang, X. Liu, & G. Fulmer (Eds.), *Chinese Science Education in the 21st Century: Policy, Practice, and Research. Contemporary Trends and Issues in Science Education* (Vol. 45, pp. 235-257). Dordrecht: Springer. Retrieved from https:// doi.org/10.1007/978-94-017-9864-8_10

Churchill, D. (2018). Transformation of Teaching through 3D Learning-Outcomes Based Curriculum Approach in Higher Education. In T.R. Dash, & M. Behere (Eds.), *Educational Access and Excellence* (pp. 66-75). Allied Publishers Private Limited: New Delhi.

CISCO (2017). Cisco Visual Networking Index: Forecast and Methodology, 2016–2021, *White Paper*. Retrieved from https://www.cisco.com/c/en/us/solutions/collateral/ service-provider/ visualnetworking-index-vni/complete-white-paper-c11-481360.html

Clariana, R., & Wallace, P. (2002). Paper–based versus computer–based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology*, *33*(5), 593-602. Retrieved from https://doi.org/10.1111/1467-8535.00294

Clark, C. (2013). *Accelerated Reader and Young People's Reading. Findings from the National Literacy Trust's 2012 Annual Literacy Survey on Reading Enjoyment, Reading Behaviour outside Class and Reading Attitudes.* London: National Literacy Trust. Retrieved from http://www.litera cytrust.org.uk/assets/0001/9353/ AR_and_young_people_s_reading.pdf.

Clark, R. C., Nguyen, F., & Sweller, J. (2011*). Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Francisco, CA: Pfeiffer.

Clark, T. (2008). We're Over-Researched Here!' Exploring Accounts of Research Fatigue within Qualitative Research Engagements. *Sociology*, *42*(5), 953-970. London: SAGE Publications Ltd. doi:10.1177/0038038508094573

Coccia, M.(2018). *What Is Science and Scientific Research? Why Do Nations Produce Scientific Research?* Working Paper CocciaLab n. 35, CNR - National Research Council of Italy (2018). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_ id=3190992

Cohen, L., Manion, L., & Morrison, K. (2017). *Research Methods in Education* (8th ed.). Abingdon, Oxon: Routledge.

Cole, M. S., Feild, H. S., & Harris, S. G. (2004). Student learning motivation and psychological hardiness: Interactive effects on students' reactions to a management class. *Academy of Management Learning & Education*, *3*(1), 64-85. Retrieved from https://doi.org/10.5465/amle.2004.12436819

Compton, H. (2013). A Historical Overview of M-learning. In Z. L. Berge, & L. Muilenburg (Eds.), *Handbook of mobile learning* (pp. 3-12). New York, NY: Routledge.

Computer History Museum (2018). http://www.computerhistory.org/timeline/computers/ Accessed 03 February 2018.

Cote, L.R., Kwak, K., Putnick, D.L., Chung, H.J., & Bornstein, M.H. (2015). The Acculturation of Parenting Cognitions: A Comparison of South Korean, Korean Immigrant, and European American Mothers. *Journal of Cross-Cultural Psychology, 46*(9), 1115-1130. Retrieved from https://doi.org/10.1177/0022022115600259

Cotton, J. (1995). *The Theory of Assessment*. *An Introduction*. Great Britain: Biddles Ltd.

Cowan, N. (1999). An Embedded-Processes Model of Working Memory. In A. Miyake, & S. Priti (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62-101). USA: Cambridge University Press.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87-185. Retrieved from https://www.cambridge.org/core/services/aop-cambridge-core/content/view/727115B70056D3A90A19EB58AD1F0C84/S0140525X01003922a.pdf/metatheory_of_storage_capacity_limits.pdf

Cowan, N. (2014). Working Memory Underpins Cognitive Development, Learning, and Education. *Educational Psychology Review*, 26, 197-223. doi:10.1007/s10648-013-9246-y

Cowan, N. (1997). The Development of Working Memory. In N. Cowan, & C. Hulme (Eds.), *The Development of Memory in Childhood* (pp. 163-199). New York, NY: Psychology Press Ltd.

Craik, K. (1943). *The Nature of Explanation*. London: Cambridge University Press.

Craik, F., & Lockhart, R.S. (1972). Levels of Processing: A Framework for Memory Research. *Journal of Verbal Learning and Verbal Behaviour*, *11*, 671-684. Retrieved from https://doi.org/10.1016/S0022-5371(72)80001-X

Credé, M., & Phillips, L. A. (2011). A Meta-Analytic Review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences*, *21*(4), 337-346. doi:10.1016/j.lindif.2011.03.002

Creswell, J.W., & Creswell, J.D. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Los Angeles: SAGE Publications Ltd.

Cromack, J. (2008). Technology and Learning-Centered Education: Research-Based Support for How the Tablet PC Embodies the Seven Principles of Good Practice in Undergraduate Education. In *2008 38th Annual Frontiers in Education Conference* (pp. T2A-1). IEEE. doi:10.1109/FIE.2008.4720288

Crookes, G. (2013). Epistemology and Ontology. *The Encyclopaedia of Applied Linguistics*. Retrieved from https://doi.org/10.1002/9781405198431.wbeal0390

Cropley, A. J., & Field, T. W. (1969). Achievement in Science and Intellectual Style. *Journal of Applied Psychology*, 53(2, Pt.1), 132. Retrieved from http://dx.doi.org/10.1037/h0027087

Crotty, M. (1998). *The Foundations of Social Research: Meaning and Perspective in the Research Process.* London, UK: SAGE Publications Ltd.

Daniel, D. B., & Woody, W. D. (2013). E-textbooks at what cost? Performance and use of electronic v. print texts. *Computers & Education*, *62*, 18-23. Retrieved from http://dx.doi.org/10.1016/j.compedu.2012.10.016

Davis, S. (2003). Observations in classrooms using a network of handheld devices. *Journal of Computer Assisted Learning*, *19*(3), 298-307. Retrieved from https://doi.org/10.1046/j.0266-4909.2003.00031.x

De Bruyckere, P., Kirschner, P. A., & Hulshof, C. D. (2016). Technology in Education: What Teachers Should Know. *American Educator*, *40*(1), 12. Retrieved from https://eric.ed.gov/?id=EJ1094203

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.

Dede, C., Clarke, J., Ketelhut, D. J., Nelson, B., & Bowman, C. (2005, April). Students' Motivation and Learning of Science in a Multi-User Virtual Environment. In *American Educational Research Association Conference* (pp. 1-8). Montreal, Canada. Retrieved from http://muve.gse.harvard.edu/rivercityproject/documents/motivation_muves_aera_2005.pdf

De Jong, T. (2010). Cognitive Load Theory, Educational Research, and Instructional Design: Some Food for Thought. *Instructional Science*, *38*(2), 105-134. doi:10.1007/s11251-009-9110-0

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, *100*(1), 223-234.

Della Sala, S., Gray, C., Baddeley, A., Alamano, N., & Wilson, L. (1999). Pattern Span: a tool for unwielding visuo-spatial memory. *Neuropsychologia*, *37*(10), 1189-1199. Retrieved from https://www.sciencedirect.com/science/article/pii/S0028393298001596

Demoulin, C. & Kolinsky, R. (2016). Does learning to read shape verbal working memory? *Psychonomic Bulletin & Review*, *23*(3), 703-722. Retrieved from https://link.springer.com/article/10.3758/s13423-015-0956-7

Dewey, J. (1929). My Pedagogic Creed. *Journal of the National Education Association, 18* (9), 291-295. Retrieved from http://www.infed.org/archives/e-texts/e-dew-pc.htm

Dewey, J. (1959). *The child and the curriculum* (No. 5). Chicago: University of Chicago Press.

Dimitrov, D. M., & Rumrill Jr, P. D. (2003). Pretest-posttest designs and measurement of change. *Work*, *20*(2), 159-165.

Dindar, M., Kabakçı Yurdakul, I., & İnan Dönmez, F. (2015). Measuring cognitive load in test items: static graphics versus animated graphics. *Journal of Computer Assisted Learning*, *31*(2), 148-161. Retrieved from https://doi.org/10.1111/jcal.12086

Dixon, R. M., Verenikina, I., Costley, D., & Pryor, S. (2015). The Use of iPADs in the Home Setting for Students with Autism Spectrum Disorders. *Journal of Special Education Technology*, *30*(4), 193-206. doi:10.1177/0162643415623023

Doharty, N. (2017). *Black History Month and Black History with Key Stage 3 Students in English Secondary Schools: A Critical Race Theory Approach.* Doctoral thesis, Keele University. Retrieved from http://eprints.keele.ac.uk/4111/

Driessen, E. W., Muijtjens, A. M., Van Tartwijk, J., & Van Der Vleuten, C. P. (2007). Web-or paper-based portfolios: is there a difference? *Medical Education*, *41*(11), 1067-1073. Retrieved from https://doi.org/10.1111/j.1365-2923.2007.02859.x

Du, N.N. (2012). Educational Research: Purpose, Quality and Effectiveness. *The Annual Review of Education, Communication, and Language Sciences*, *9*, 2-62. Retrieved from https://research.ncl.ac.uk/ARECLS/volume_9/ngocdu_vol9.pdf

Duchesne, S., & McMaugh, A. (2018). *Educational psychology for learning and teaching.* South Melbourne: Cengage Learning Australia.

Ebbinghaus, H. (2013). Memory: A contribution to experimental psychology. *Annals of neurosciences*, *20*(4), 155-156. (Original work published 1923).

Ecke, P. (2004). Language attrition and theories of forgetting: A cross-disciplinary review. *International Journal of Bilingualism*, *8*(3), 321-354. Retrieved from http://journals. sagepub.com/doi/pdf/10.1177/13670069040080030901

Ehri, L., Nunes, S., Willows, D., Schuster, B., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic Awareness Instruction Helps Children Learn to Read: Evidence from the National Reading Panel's Meta-Analysis. *Reading Research Quarterly*, *36*(3), 250-287.

Eimer, M. (2015). Visual Working Memory and Attentional Object Selection. In P. Jolicoeur, C. Lefebvre, & J. Martinez-Trujillo (Eds.), *Mechanisms of Sensory Working Memory: Attention and Performance XXV* (pp. 89-104). USA: Elsevier.

Eisner, E.W. (2017). *The Enlightened Eye: Qualitative Inquiry and the Enhancement of Educational Practice*. New York, NY: Teachers College Press.

Ellis, K. (2008). Cyber Charter Schools: Evolution, Issues, and Opportunities in Funding and Localized Oversight. *Educational Horizons*, *86*(3), 142-152.

El Zein, F., Gevarter, C., Bryant, B., Son, S. H., Bryant, D., Kim, M., & Solis, M. (2016). A Comparison between iPad-Assisted and Teacher-Directed Reading Instruction for Students with Autism Spectrum Disorder (ASD). *Journal of Developmental and Physical Disabilities, 28*(2), 195-215. Retrieved from https://doi.org/10.1007/s1088

Emerson, L., & MacKay, B. (2011). A comparison between paper-based and online learning in higher education. *British Journal of Educational Technology*, *42*(5), 727-735. Retrieved from https://doi.org/10.1111/j.1467-8535.2010.01081.x

Enriquez, A.G. (2010). Enhancing Student Performance Using Tablet Computers. *College Teaching, 58,* 77-84. Retrieved from https://doi.org/10.1080/87567550903263859

Epstein, M. L., Phillips, W. D., & Johnson, S. J. (1975). Recall of related and unrelated word pairs as a function of processing level. *Journal of Experimental Psychology: Human Learning and Memory, 1*(2), 149-152. Retrieved from http://dx.doi.org/10.1037/0278-7393.1.2.149

Evans, S., Beauchamp, G., & John, V. (2014). 'Learners' experience and perceptions of informal learning in Key Stage 3 music lessons: a collective case study, exploring the implementation of Musical Futures in three secondary school in Wales', *Music Education Research, 17* (1), 1-16. Retrieved from https://repository.cardiffmet .ac.uk/handle/10369/7663

Eyre, D. (2016). *High Performance Learning. How to Become a World Class School*. Abingdon, Oxon: Routledge.

Falloon, G. (2013). Young students using iPads: App design and content influences on their learning pathways. *Computers & Education*, *68*, 505-521. Retrieved from http://dx.doi.org/10.1016/j.compedu.2013.06.006

Faraway, J. J. (2002). Practical regression and ANOVA using R. Retrieved from: http: //cran.r-project.org/doc/contrib/Faraway-PRA.pdf

Faraway, J. J. (2016). *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models* (Vol. 124). CRC press.

Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature*, *425*(6958), 614-616. doi:10.1038/ nature01951

Field, A. (2014). *Discovering statistics using SPSS*. London, UK: SAGE Publications Ltd.

Fiorella, L., & Mayer, R. E. (2012). Paper-based aids for learning with a computer-based game. *Journal of Educational Psychology*, *104*(4), 1074.

Fischer, K. W., Daniel, D. B., Immordino-Yang, M. H., Stern, E., Battro, A., & Koizumi, H. (2007). Why Mind, Brain, and Education? Why Now? *Mind, Brain, and Education, 1*(1), 1-2. Retrieved from https://doi.org/10.1111/j.1751-228X.2007.00006.x

Fitó-Bertran, A., Hernández-Lara, A.B., & Serradell-López, E. (2013). Comparing student competences in a face-to-face and online business game. *Computers in Human Behavior, 30*, 452-459. Retrieved from http://dx.doi.org/10.1016/j.chb.2013.06.023

Fleck, R., & Fitzpatrick, G. (2010). Reflecting on reflection: framing a design landscape. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction* (pp. 216-223). ACM. doi:10.1145/1952222.1952269

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological bulletin*, *101*(2), 171-191.

Fojtik, R. (2015). Ebooks and mobile devices in education. *Procedia-Social and Behavioral Sciences*, *182*, 742-745. doi:10.1016/j.sbspro.2015.04.824

Foulger, T. S., Amrein-Beardsley, A., & Toth, M. J. (2011). Students' Roles in Exposing Growing Pains: Using the "Dean's Concerns" to Refine Hybrid Instruction. *International Journal of Teaching and Learning in Higher Education*, *23*(2), 150-165.

Furberg, A. (2009). Socio-cultural aspects of prompting student reflection in Web-based inquiry learning environments. *Journal of Computer Assisted Learning*, *25*(4), 397-409. Retrieved from https://doi.org/10.1111/j.1365-2729.2009.00320.x

Gabrielle, D. M. (2003). *The Effects of Technology-Mediated Instructional Strategies on Motivation, Performance, and Self-Directed Learning*. Thesis from State Florida University, DigiNole: FSU's digital repository, http://diginole.lib.fsu.edu/islandora/object/fsu%3A182576

Gagné, R. M. (2013). *Instructional technology: foundations*. Abingdon, Oxon: Routledge.

Ganley, C. M., & Vasilyeva, M. (2014). The role of anxiety and working memory in gender differences in mathematics. *Journal of Educational Psychology*, *106*(1), 105.

García, T., & Pintrich, P. R. (1995). Assessing Students' Motivation and Learning Strategies: The Motivated Strategies for Learning Questionnaire. Retrieved from https://files.eric.ed.gov/fulltext/ED383770.pdf

Gardner, H. (1983, 2011). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gathercole, S. E., Hitch, G. E., & Service, E. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology*, *33*(6), 966-979.

Gathercole, S. E., Lamont, E., & Alloway, T. P. (2006). Working memory in the classroom. In S. J. Pickering (Ed.), *Working memory and education* (pp. 219-240). London, UK: Academic Press.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental psychology*, *40*(2), 177.

Gathercole, S. E., Woolgar, F., Kievit, R. A., Astle, D., Manly, T., Holmes, J., & CALM Team. (2016). How Common are WM Deficits in Children with Difficulties in Reading and Mathematics*? Journal of Applied Research in Memory and Cognition, 5*(4), 384-394. Retrieved from https://doi.org/10.1016/j.jarmac. 2016.07.013

Ge, Y.P., Unsworth, L., Wang, K.H., Chang, H.P. (2018). Image Design for Enhancing Science Learning: Helping Students Build Taxonomic Meanings with Salient Tree Structure Images. In: K.S. Tang, K. Danielsson (Eds.), *Global Developments in Literacy Research for Science Education.* Springer, Cham

Gevins, A., & Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex, 10*(9), 829-839.

Giguère, G. & Love, B.C. (2013). Limits in decision making arise from limits in memory retrieval. *PNAS, 110* (19) 7613-7618. Retrieved from https://doi.org/10.1073/pnas. 1219674110

Gillmor, S. C., Poggio, J., & Embretson, S. (2015). Effects of reducing the cognitive load of mathematics test items on student performance. *Numeracy*, *8*(1), 4. Retrieved from http://dx.doi.org/10.5038/1936-4660.8.1.4

Gillon, G., & McNeill, B. (2009). Phonological awareness - Effective practices in assessment and intervention. *ACQuiring knowledge in speech, language and hearing*, *11*(2), 72-76.

Gómez-Pablos, V. B., del Pozo, M. M., & Muñoz-Repiso, A. G. V. (2017). Project-based learning (PBL) through the incorporation of digital technologies: An evaluation based on the experience of serving teachers. *Computers in Human Behavior*, *68*, 501-512. Retrieved from https://doi.org/10.1016/j.chb.2016.11.056

Goodchild, M.F., & Janelle, D.G. (2010). Towards critical spatial thinking in the social sciences and humanities, *GeoJournal, 75*, 3-13. Retrieved from https://link.springer. com/content/pdf/10.1007%2Fs10708-010-9340-3.pdf

Gorin, J. S., & Embretson, S. E. (2006). Item diffficulty modeling of paragraph comprehension items. *Applied Psychological Measurement*, *30*(5), 394-411. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/0146621606288554

Gottfredson, L.S. (1997). Mainstream Science on Intelligence: An Editorial with 52 Signatories, History, and Bibliography. *Intelligence*, *24*(1), 13-23.

Graf, S., & Kinshuk, (2008). *Handbook of Research on Educational Communications and Technology* (3rd ed.). In J. M. Spector, M. D. Merrill, & M. P. Driscoll (Eds.). New York: Lawrence Erlbaum Associates, Taylor &Francis Group

Gravetter, F. J., & Forzano, L. A. B. (2018). *Research methods for the behavioral sciences.* Cengage Learning.

Graves, M. F. (1986). Chapter 2: Vocabulary Learning and Instruction. *Review of research in education, 13*(1), 49-89. Retrieved from https://doi.org/10.3102/ 0091732X 013001049

Greenfield, P. M. (2014). *Mind and media: The effects of television, video games, and computers*. New York, NY: Psychology Press.

Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology, 61*(1), 29-48.

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters.* Advanced Analytics, LLC.

Hall, P. (2003). Aligning Ontology and Methodology in Comparative Research. In J. Mahoney, & D. Rueschemeyer (Eds.), *Comparative Historical Analysis in the Social* Sciences (pp. 373-406). New York: Cambridge University Press.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology, 8*(1), 23.

Halpern, D. F. (2013). *Sex differences in cognitive abilities* (4th ed.). New York, NY: Psychology press.

Hammersley, M., & Traianou, A. (2012). *Ethics in Qualitative Research: Controversies and Contexts*. London: SAGE Publications Ltd.

Handlin, O. (1959). *John Dewey's Challenge to Education: Historical Perspectives on the Cultural Context*. New York: Harper.

Hanham, J., Leahy, W., & Sweller, J. (2017). Cognitive load theory, element interactivity, and the testing and reverse testing effects. *Applied Cognitive Psychology, 31*(3), 265-280.

Hanson, D. L. (1967). Influence of the Hawthorne Effect upon Physical Education Research. *Research Quarterly. American Association for Health, Physical Education and Recreation*, *38*(4), 723-724. doi:10.1080/10671188.1967.10616519

Harasim, L. (2012). *Learning Theory and Online Technologies.* New York: Routledge.

Harness, A., Jacot, L., Scherf, S., White, A., & Warnick, J. E. (2008). Sex differences in working memory. *Psychological reports*, *103*(1), 214-218. Retrieved from http://journals.sagepub.com/doi/abs/10.2466/pr0.103.1.214-218

Harris, J., Mishra, P., & Koehler, M. (2009). Teachers' technological pedagogical content knowledge and learning activity types: Curriculum-based technology integration reframed*. Journal of Research on Technology in Education*, *41*(4), 393-416. Retrieved from https://doi.org/10.1080/15391523.2009.10782536

Hart, J. W., Stasson, M. F., Mahoney, J. M., & Story, P. (2007). The Big Five and achievement motivation: Exploring the relationship between personality and a two-factor model of motivation. *Individual Differences Research*, *5*(4), 267-274.

Hartnell-Young, E., & Heym, N. (2008). *How mobile phones help learning in secondary schools*. Learning Sciences Research Institute, University of Nottingham.

Hawlitschek, A., & Joeckel, S. (2017). Increasing the effectiveness of digital educational games: The effects of a learning instruction on students' learning, motivation and cognitive load. *Computers in Human Behavior*, *72*, 79-86. Retrieved from https://doi.org/10.1016/j.chb.2017.01.040

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-based nursing*, ebnurs-2015.

Healy, M., & Perry, C. (2000). Comprehensive criteria to judge validity and reliability of qualitative research within the realism paradigm. *Qualitative market research: An international journal, 3*(3), 118-126. Retrieved from https://doi.org/10.1108/13522750010333861

Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory.* London: Lawrence Erlbaum Associates, Publishers. (Original work published 1949).

Herrington, J., Mantei, J., Herrington, A., Olney, I. W., & Ferry, B. (2008). New technologies, new pedagogies: Mobile technologies and new ways of teaching and learning. In R. Atkinson, & C. McBeath (Eds.), *Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education*, *2008* (pp. 419-427). Melbourne, Australia: Deakin University.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: The reshaping of American life by differences in intelligence*. New York: The Free Press. Retrieved from https://lesacreduprintemps19.files.wordpress.com/2012/11/the-bell-curve.pdf

Hickey, D.T. (1997). Motivation and contemporary socio-constructivist instructional perspectives, *Educational Psychologist, 32*(3), 175-193. Retrieved from https://doi.org/10.1207/s15326985ep3203_3

Hill, B. D., Foster, J. D., Sofko, C., Elliott, E. M., & Shelton, J. T. (2016). The interaction of ability and motivation: Average working memory is required for Need for Cognition to positively benefit intelligence and the effect increases with ability. *Personality and Individual Differences, 98,* 225-228. Retrieved from https://doi.org/10.1016/j.paid.2016.04.043

Hodge, S. (2018). Standardised curriculum and hermeneutics: the case of Australian vocational educators. *Journal of Curriculum Studies, 50*(1), 38-55. Retrieved from https://doi.org/10.1080/00220272.2016.1278042

Hoerig, D. C., David, A. S., & D'Amato, R. C. (2002). Evaluating the relation between memory and intelligence in children with learning disabilities. *Psychological reports*, *91*(3_suppl), 1169-1173. Retrieved from https://doi.org/10.2466/pr0.2002.91.3f.1169

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and instruction*, *17*(6), 722-738. Retrieved from https://doi.org/10.1016/j.learninstruc.2007.09.013

Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. USA: CRC Press.

Homer, B.D., Plass, J. L., & Blake, L. (2008). The effects of video on cognitive load and social presence in multimedia-learning. *Computers in Human Behavior, 24*(3), 786-797. Retrieved from https://doi.org/10.1016/j.chb.2007.02.009

Hunt, K. J., Shlomo, N., & Addington-Hall, J. (2013). Participant recruitment in sensitive surveys: a comparative trial of ‘opt in’versus ‘opt out’approaches. *BMC medical research methodology, 13*(3), 1-8. Retrieved from https://bmcmedresmethodol. biomedcentral.com/track/pdf/10.1186/1471-2288-13-3

Hyde, J. S. (2007). New directions in the study of gender similarities and differences. *Current Directions in Psychological Science*, *16*(5), 259-263. Retrieved from https://doi.org/10.1111/j.1467-8721.2007.00516.x

Iten, N., & Petko, D. (2016). Learning with serious games: Is fun playing the game a predictor of learning success? *British Journal of Educational Technology, 47*(1), 151-163. Retrieved from https://doi.org/10.1111/bjet.12226

Jabr, F. (2013, April). The reading brain in the digital age: The science of paper versus screens. *Scientific American*. Retrieved from: http://www. scientificamerican.com/article/reading-paper-screens/.

Jang, Y. S. (2000). The worldwide founding of ministries of science and technology, 1950–1990. *Sociological Perspectives*, *43*(2), 247-270. Retrieved from http://journals.sagepub.com/doi/pdf/10.2307/1389796

John Hopkins Medicine. (2018, August 14). Anatomy and Function of the Heart Valves. Retrieved from https://www.hopkinsmedicine.org/healthlibrary/conditions/ cardiovascular_diseases/anatomy_and_function_of_the_heart_valves_90,P03059

Johnson, R.B., Onwuegbuzie, A.J., & Turner, L.A. (2007). Towards a Definition of Mixed Methods Research. *Journal of Mixed Methods Research, 1*(2), 112-133. Retrieved from http://journals.sagepub.com/doi/pdf/10.1177/1558689806298224

Johnson, S. (2006). The sleeper curve. In *Everything bad is good for you: How today’s popular culture is actually making us smarter* (pp. 1-14). New York, NY: Penguin Group.

Jonassen, D. H. (2004a). *Learning to solve problems: An instructional design guide* (Vol. 6). San Francisco, CA: Pfeiffer.

Jonassen, D. H. (2004b). *Handbook of research on educational communications and technology* (2nd ed.). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Kabali, H. K., Irigoyen, M. M., Nunez-Davis, R., Budacki, J. G., Mohanty, S. H., Leister, K. P., & Bonner, R. L. (2015). Exposure and use of mobile media devices by young children. *Pediatrics*, *136*(6), 1044-1050. Retrieved from http://pediatrics. aappublications.org/content/136/6/1044

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing Split-attention and Redundancy in Multimedia Instruction. *Applied Cognitive Psychology*, *13*, 351-371.

Kalyuga, S., & Liu, T. C. (2015). Guest Editorial: Managing Cognitive Load in Technology-Based Learning Environments. *Educational Technology & Society*, *18*(4), 1-8.

Kaptelinin, V., & Nardi, B. A. (2006). *Acting with technology: Activity theory and interaction design*. Cambridge, MA: MIT Press.

Karunanayaka, S. P., Naidu, S., Rajendra, J.C.N., & Ratnayake, H.U.W. (2017). Designing Reflective Practice in the Context of OER-based e-Learning. *Journal of Learning for Development*, *4*(2), 143-160. Retrieved from http://oasis.col.org/handle/11599/2777

Katz, Y. J., & Yablon, Y. B. (2002). Who is afraid of university Internet courses? *Educational Media International*, *39*(1), 69-73. Retrieved from https://doi.org/ 10.1080/09523980210131132

Keengwe, J., Schnellert, G., & Jonas, D. (2014). Mobile phones in education: Challenges and opportunities for learning. *Education and Information Technologies*, *19*(2), 441-450. Retrieved from https://doi.org/10.1007/s10639-012-9235-7

Khoshsima, H., Hosseini, M., & Hashemi Toroujeni, S. M. (2017). Cross-Mode Comparability of Computer-Based Testing (CBT) Versus Paper-Pencil Based Testing (PPT): An Investigation of Testing Administration Mode among Iranian Intermediate EFL Learners. *English Language Teaching*, *10*(2), 23-31. Retrieved from http://dx.doi.org/10.5539/elt.v10n2p23.

Kim, Y. S., Al Otaiba, S., Wanzek, J., & Gatlin, B. (2015). Toward an understanding of dimensions, predictors, and the gender gap in written composition. *Journal of educational psychology*, *107*(1), 79. Retrieved from http://dx.doi.org/10.1037/ a0037210

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological review*, *85*(5), 363. Retrieved from http://dx.doi.org/10.1037/0033-295X.85.5.363

Kori, K., Pedaste, M., Leijen, Ä., & Mäeots, M. (2014). Supporting reflection in technology-enhanced learning. *Educational Research Review*, *11*, 45-55. Retrieved from https://doi.org/10.1016/j.edurev.2013.11.003

Krapohl, E., Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J., … Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proc. Natl. Acad. Sci. U.S.A.*, 111, 15273–15278. doi:10.1073/pnas.1408777111pmid:25288728

Kress, G., Sharon, B., & Bassan, D. (1981). *Work Motivation. Theory and Practice.* Virginia: U.S. Army Research Institute for the Behavioural and Social Sciences.

Krishnan, S., Watkins, K., & Bishop, D. V. M. (2017). The effect of recall, reproduction, and restudy on word learning: a pre-registered study. *BMC Psychology,* 5(28), 2-14. Retrieved from https://bmcpsychology.biomedcentral.com/track/pdf/10.1186/s40359-017-0198-8

Labrie, A. M., Lemanowski, V. L., Smiley, B. A., Yezierski, E. J., Baker, D. R., & Anderson-Rowland, M. R. (2003). A new look at gender equity professional development for secondary science/mathematics teachers and counselors. In *ASEE Annual Conference Proceedings* (pp. 3041-3059).

Lecerf, T., & de Ribaupierre, A. (2005). Recognition in a visuospatial memory task: The effect of presentation. *European Journal of Cognitive Psychology*, *17*(1), 47-75. Retrieved from https://doi.org/10.1080/09541440340000420

Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, *3*, 18-23. Retrieved from https://doi.org/10.1016/j.mfglet.2014.12.001

Lee, J., & Bednarz, R. (2012). Components of spatial thinking: Evidence from a spatial thinking ability test. *Journal of Geography, 111*(1), 15–26. Retrieved from https://doi.org/10.1080/00221341.2011.583262

Lehrer, K. (2018). *Theory of Knowledge* (2nd ed.). New York: Routledge.

Leijen, Ä., Valtna, K., Leijen, D. A., & Pedaste, M. (2012). How to determine the quality of students' reflections? *Studies in Higher Education*, *37*(2), 203-217. Retrieved from https://doi.org/10.1080/03075079.2010.504814

Leinonen, T., Keune, A., Veermans, M., & Toikkanen, T. (2016). Mobile apps for reflection in learning: A design research in K-12 education, *British Journal of Educational Technology*, *47*(1), 184–202. Retrieved from https://doi.org/10.1111/bjet.12224

LeVine, R. (2017). Culture, Behavior, and Personality: An Introduction to the Comparative Study of Psychosocial Adaptation (2nd ed.). New York: Routledge.

Li, Q. (2007). Student and teacher views about technology: A tale of two cities? *Journal of research on Technology in Education, 39*(4), 377-397. doi: 10.1080/15391523.2007. 10782488

Lin, L., Atkinson, R. K., Savenye, W. C., & Nelson, B. C. (2016). Effects of visual cues and self-explanation prompts: empirical evidence in a multimedia environment. *Interactive Learning Environments, 24*(4), 799-813. Retrieved from https://doi.org/10.1080/10494820.2014.924531

Lin, M. H., Chen, H. C., & Liu, K. S. (2017). A study of the effects of digital learning on learning motivation and learning outcome. *Eurasia Journal of Mathematics, Science and Technology Education, 13*(7), 3553-3564.

Liu, Y., & Shrum, L. J. (2009). A dual-process model of interactivity effects. *Journal of Advertising, 38*(2), 53-68.

Lockhart, R.S. (2002). Levels of Processing, transfer-appropriate processing, and the concept of robust encoding. *Memory, 10,* 397-403. doi: 10.1080/09658210244000225

Lockhart, R.S., & Craik, F.I.M. (1990). Levels of Processing: A retrospective analysis of a framework for memory research. *Canadian Journal of Psychology, 44*, 87-112. Retrieved from http://dx.doi.org/10.1037/h0084237

Logie, R.H. (1995). *Visuo-spatial Working Memory*. Hove: Lawrence Erlbaum Associates Ltd. Retrieved from http://old.nbu.bg/cogs/events/2002/materials/Markus/Visuo-spatial%20Working%20Memory.pdf

Looi, C. K., Seow, P., Zhang, B. H., So, H. J., Chen, W., & Wong, L. H. (2010). Leveraging mobile technology for sustainable seamless learning: A research agenda. *British Journal of Educational Technology, 41*(2), 154–169. Retrieved from https://doi.org/10.1111/j.1467-8535.2008.00912.x

Looi, C. K., Wong, L. H., So, H. J., Seow, P., Toh, Y., Chen, W., Zhang, B., Norris, C., & Soloway, E. (2009). Anatomy of a mobilized lesson: Learning my way. *Computers & Education*, *53*(4), 1120-1132. doi:10.1016/j.compedu.2009.05.021

Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and instruction*, *13*(2), 157-176. doi:10.1016/S0959-4752(02)00018-X

Lowyck, J. (2008). *Handbook of Research on Educational Communications and Technology* (3rd ed.). In J. M. Spector, M. D. Merrill, & M. P. Driscoll (Eds.). New York: Lawrence Erlbaum Associates, Taylor &Francis Group

Lucid Recall (2013*). Computerised Assessment of Working Memory and Processing Speed for 7 to 16 Years, Administrator's Manual*. East Yorkshire, UK: Lucid Research Ltd. Retrieved from http://lucid-research.net/documents/manuals/manual_rc.pdf

Lynch, J., & Redpath, T. (2014). 'Smart' Technologies in Early Years Literacy Education: A Meta-Narrative of Paradigmatic Tensions in iPad Use in an Australian Preparatory Classroom. *Journal of Early Childhood Literacy*, *14*(2), 147-174. Retrieved from http://journals.sagepub.com/doi/pdf/10.1177/1468798412453150

Machin, S., & Vignoles, A. (2005). Education Policy in the UK. *CESifo DICE Report, ISSN 1613-6373, 3*(4), 64-74. Retrieved from https://www.econstor.eu/bitstream/10419/166864/1/ifo-dice-report-v03-y2005-i4-p64-74.pdf

Majerus, S., Van der Linden, M., Mulder, L., Meulemans, T., & Peters, F. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language*, *51*(2), 297-306. doi:10.1016/j.jml.2004.05.002

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*(3), 252-271. Retrieved from http://dx.doi.org/10.1037/0033-295X.87.3.252

Mangen, A. (2017). Textual reading on paper and screens. *Information Design: Research and Practice*, 275-290. Abingdon: Routledge.

Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research, 58*, 61-68. Retrieved from https://doi.org/10.1016/j.ijer.2012.12.002

Martens, R., Gulikers, J., & Bastiaens, T. (2004). The impact of intrinsic motivation on e-learning in authentic computer tasks. *Journal of Computer Assisted Learning*, *20*(5), 368-376. Retrieved from https://doi.org/10.1111/j.1365-2729.2004.00096.x

Martinez, B. (2010). The influence of computer self-efficacy, self-regulation, and engagement on achievement (Doctoral dissertation). University of Southern California. Retrieved from B. Martinez via brandodm@usc.edu.

Mathy, F., Chekaf, M., & Cowan, N. (2018). Simple and Complex Working Memory Tasks Allow Similar Benefits of Information Compression. *Journal of Cognition, 1*(1). Doi: http://doi.org/10.5334/joc.31

Maxwell, N. P. (2018). *Modeling Memory: Exploring the Relationship between Word Overlap and Single Word Norms When Predicting Judgments and Recall*. MSU Graduate Master’s Thesis, BearWorks Institutional Repository. Retrieved from https://bearworks.missouristate.edu/cgi/viewcontent.cgi?article=4272&context=theses

Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R.E. (2008). Applying the Science of Learning: Evidence-Based Principles for the Design of Multimedia Instruction. *American Psychologist*, November, 760-769. Retrieved from https://www.instructionaldesign.institute/uploads/3/1/2/2/31221959/richard_mayer_article_applying_the_science_of_learning.pdf

Mayer, R. E. (2009). *Multimedia Learning*. New York, NY: Cambridge University Press.

Mayer, R. E. (2014). Incorporating motivation into multimedia learning. *Learning and Instruction, 29*, 171-173. Retrieved from http://dx.doi.org/10.1016/j.learninstruc.2013.04.003

Mayer, R. E., DeLeeuw, K. E., & Ayres, P. (2007). Creating retroactive and proactive interference in multimedia learning. *Applied Cognitive Psychology*, *21*(6), 795-809. Retrieved from https://doi.org/10.1002/acp.1350

Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction, 33*, 12-18. Retrieved from https://doi.org/10.1016/j.learninstruc.2014.02.004

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational psychology review, 14*(1), 87-99. Retrieved from https://link.springer.com/content/pdf/10.1023%2FA%3A1013184611077.pdf

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, *38*(1), 43-52. Retrieved from https://doi.org/10.1207/S15326985EP3801_6

McGinnis, J. R., & Tippins, D. J. (2001). Gender: Contributions by Wendy Giles, Peter Leach. In J. Wallace, & W. Louden (Eds.), *Dilemmas of Science Teaching: Perspectives on Problems of Practice* (pp. 59-72). New York, NY: RoutledgeFalmer.

McMahon, D., Wright, R., Cihak, D. F., Moore, T. C., & Lamb, R. (2016). Podcasts on Mobile Devices as a Read-Aloud Testing Accommodation in Middle School Science Assessment. *Journal of Science Education and Technology*, *25*(2), 263-273. Retrieved from https://doi.org/10.1007/s10956-015-9591-3

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, DC: US Department of Education. Retrieved from http://repository.alt.ac.uk/629/1/US_DepEdu_Final_report_2009.pdf

Mehlenbacher, B., Miller, C. R., Covington, D., & Larsen, J. S. (2000). Active and interactive learning online: A comparison of Web-based and conventional writing classes. *IEEE Transactions on Professional Communication, 43*(2), 166-184. doi:10.1109/47.843644

Melhuish, K., & Falloon, G. (2010). Looking to the future: M-learning with the iPad. *Computers in New Zealand Schools: Learning, Leading, Technology, 22*(3), 1-16. Retrieved from https://researchcommons.waikato.ac.nz/handle/10289/5050

Mercer, J. (2007). The challenges of insider research in educational institutions: wielding a double-edged sword and resolving delicate dilemmas. *Oxford Review of Education*, *33*(1), 1-17. doi:10.1080/03054980601094651

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, *70*, 29-40.

Meyer, S., & Ward, P. (2014). 'How to' use social theory within and throughout qualitative research in healthcare contexts. *Sociology Compass*, *8*(5), 525-539. Retrieved from https://doi.org/10.1111/soc4.12155

Miller, G.A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81-97.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of abnormal psychology*, *110*(1), 40-48.

Mingers, J., Mutch, A., & Willcocks, L. (2013). Introduction [special issue: Critical realism in information systems research]. *MIS quarterly*, *37*(3), 795-802. Retrieved from http://irep.ntu.ac.uk/id/eprint/20232/1/216160_300.pdf

Mir, R., & Watson, A. (2001). Critical realism and constructivism in strategy research: toward a synthesis. *Strategic management journal*, *22*(12), 1169-1173. Retrieved from https://doi.org/10.1002/smj.200

Mirza, A., & Hussain, N. (2018). Performing Below the Targeted Level: An Investigation into KS3 Pupils' Attitudes Towards Mathematics. *Journal of Education and Educational Development. 5*(1), 8-24. Retrieved from http://www.journals.iobmresearch.com/index.php/JoEED/article/view/1390/1

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers college record*, *108*(6), 1017-1053. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.523.3855&rep=rep1&type=pdf

Mistler-Jackson, M., & Songer, N. B. (2000). Student motivation and Internet technology: Are students empowered to learn science? *Journal of Research in Science Teaching*, *37*(5), 459-479. Retrieved from https://pdfs.semanticscholar.org/af39/2ec50009dfdfd0786b16bad052a6c2ad81ce.pdf

Mook, D. G. (1983). In defense of external invalidity. *American psychologist*, 38(4), 379.

Moon, J. A. (2013). *Reflection in learning and professional development: Theory and practice.* Abingdon, Oxon: RoutledgeFalmer.

Moreno, R. (2006). Learning in High-Tech and Multimedia Environments**.** *Current Directions in Psychological Science*, *15*(2), 63-67. Retrieved from http://journals.sagepub.com/doi/pdf/10.1111/j.0963-7214.2006.00408.x

Moreno, R. (2007). Optimising learning from animations by minimising cognitive load: Cognitive and effective consequences of signalling and segmentation methods. *Applied cognitive psychology*, *21*(6), 765-781. Retrieved from https://doi.org/ 10.1002/acp.1348

Moreno, R. (2010). Cognitive load theory: More food for thought. *Instructional Science*, *38*(2), 135-141. Retrieved from https://doi.org/10.1007/s11251-009-9122-9

Moreno, R., & Mayer, R.E. (2010). Techniques that Increase Generative Processing. In J. L. Plaas, R. Moreno, & R. Brünken (Eds.), *Cognitive Load Theory* (pp. 153-178). New York, NY: Cambridge University Press. Retrieved from doi: 10.1017/CBO 9780511844744.010

Moreno, R., & Valdez, A. (2005). Cognitive load and learning effects of having students organize pictures and words in multimedia environments: The role of student interactivity and feedback. *Educational Technology Research and Development*, *53*(3), 35-45. Retrieved from https://doi.org/10.1007/BF02504796

Morgan, G., & Smircich, L. (1980). The Case for Qualitative Research. *Academy of Management, 5*(4), 491-500.

Motiwalla, L. F. (2007). Mobile learning: A framework and evaluation. *Computers & Education*, *49*(3), 581-596. Retrieved from z

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, *87*(2), 319-334. Retrieved from http://dx.doi.org/10.1037/0022-0663.87.2.319

Muller, D. A., Sharma, M. D., & Reimann, P. (2008). Raising cognitive load with linear multimedia to promote conceptual change. *Science Education*, *92*(2), 278-296. Retrieved from https://doi.org/10.1002/sce.20244

Mwanza-Simwami, D. (2013). Activity theory and educational technology design. In R. Luckin, S. Puntambekar, P. Goodyear, B. L. Grabowski, J. Underwood, & N. Winters (Eds.), *Handbook of Design in Educational Technology* (pp. 176–188). Abingdon, Oxon: Routledge.

Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. In *Flow and the foundations of positive psychology* (pp. 239-263). Dordrecht: Springer.

Nardi, B. A. (Ed.). (1996). *Context and consciousness: activity theory and human-computer interaction.* Cambridge, MA: MIT Press.

National Research Council. (2006). *Learning to Think Spatially*. Washington, DC: The National Academies Press. Retrieved from https://doi.org/10.17226/11019.

Nedungadi, P., & Raman, R. (2012). A new approach to personalization: integrating e-learning and m-learning. *Educational Technology Research and Development*, *60*(4), 659-678. Retrieved from https://doi.org/10.1007/s11423-012-9250-9

Niccoli, A. M. (2015). The Effects of Reading Mode on Recall and Comprehension. *NERA Conference Proceedings 2014,* 3-10. Retrieved from https://opencommons.uconn.edu/cgi/viewcontent.cgi?article=1004&context=nera_2014

Njenga, J. K., & Fourie, L. C. H. (2010). The myths about e-learning in higher education. *British journal of educational technology*, *41*(2), 199-212. Retrieved from https://doi.org/10.1111/j.1467-8535.2008.00910.x

Norman, D. A. (2013). *The design of everyday things: Revised and expanded edition.* New York, NY: Basic books. Retrieved from http://www.nixdell.com/classes/HCI-and-Design-Spring-2017/The-Design-of-Everyday-Things-Revised-and-Expanded-Edition.pdf

Noddings, N. (2018). *Philosophy of Education* (4th ed.). New York: Routledge.

Noyes, J., Garland, K., & Robbins, L. (2004). Paper-based versus computer-based assessment: is workload another test mode effect*? British Journal of Educational Technology*, *35*(1), 111-113. Retrieved from https://doi.org/10.1111/j.1467-8535.2004.00373.x

Olson, M. H. & Hergenhahn, B. R. (2016). *An introduction to theories of learning* (9th ed.). New York, NY: Routledge.

Owens, J., & Adolescent Sleep Working Group (2014). Insufficient sleep in adolescents and young adults: an update on causes and consequences. *Pediatrics*, *134*(3), e921-e932. Retrieved from http://pediatrics.aappublications.org/content/pediatrics/134/3/e921.full.pdf

Oxford, E. L., & Amerstorfer, C. M. (2018) (Eds.). *Language Learning Strategies and Individual Learner Characteristics. Situating Strategy Use in Diverse Contexts.* London: Bloomsbury.

Paas, F., Van Gog, T., & Sweller, J. (2010). Cognitive Load Theory: New Conceptualizations, Specifications, and Integrated Research Perspectives. *Educational Psychology Review*, *22*(2), 115-121. Retrieved from https://doi.org/10.1007/s10648-010-9133-8

Paivio, A. (1990). *Mental Representations: A Dual Coding Approach.* New York, NY: Oxford University Press.

Paivio, A. (2007, 2014). *Mind and Its Evolution: A Dual Coding Theoretical Approach*. New York, NY: Psychology Press.

Papastergiou, M. (2009). Digital Game-Based Learning in High School Computer Science Education: Impact on Educational Effectiveness and Student Motivation. *Computers & Education*, *52*(1), 1-12. Retrieved from https://doi.org/10.1016/j.compedu.2008.06.004

Park, B., Plass, J. L., & Brünken, R. (2014). Cognitive and Affective Processes in Multimedia Learning. *Learning and Instruction*, *29*, 125-127. Retrieved from https://doi.org/10.1016/j.learninstruc.2013.05.005

Passey, D., & Zozimo, J. (2016). Developing mobile learning practices through teacher education: outcomes of the MLEARN pilot. *Interactive Technology and Smart Education*, *13*(1), 36-51. Retrieved from https://doi.org/10.1108/ITSE-01-2016-0002

Pellegrino, J. W. (2018): Sciences of learning and development: Some thoughts from the learning sciences, *Applied Developmental Science*. Retrieved from DOI: 10.1080/10888691.2017.1421427

Penney, C. G. (1974). Dichotic listening and sequential associations in auditory short-term memory. *Journal of Experimental Psychology*, *102*(1), 142-150. Retrieved from http://dx.doi.org/10.1037/h0035702

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, *17*(4), 398-422. Retrieved from https://link.springer.com/ content/pdf/10.3758%2FBF03202613.pdf

Pérez-Sanagustín, M., Nussbaum, M., Hilliger, I., Alario-Hoyos, C., Heller, R. S., Twining, P., & Tsai, C. C. (2017). Research on ICT in K-12 schools-A review of experimental and survey-based studies in computers & education 2011 to 2015. *Computers & Education*, *104*(C), A1-A15. Doi:10.1016/j.compedu.2016.09.006

Phan, H. P. (2014). Self-efficacy, Reflection, and Achievement: A Short-Term Longitudinal Examination. *The Journal of Educational Research*, 107(2), 90-102. Retrieved from https://doi.org/10.1080/00220671.2012.753860

Phillips, T., Porticella, N., Constas, M., & Bonney, R. (2018). A Framework for Articulating and Measuring Individual Learning Outcomes from Participation in Citizen Science. *Citizen Science: Theory and Practice, 3*(2).

Piaget, J. (1967). *Six psychological studies* (Elkind, Ed.). New York: Random House.

Pintrich, P. R. (1999). The Role of Motivation in Promoting and Sustaining Self-Regulated Learning. *International Journal of Educational Research, 31*(6), 459-470. Retrieved from https://doi.org/10.1016/S0883-0355(99)00015-4

Pintrich, P. R. (2003). Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts. *Journal of Educational Psychology, 95*(4), 667-686.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). A Manual for the Use of the Motivated Strategies for Learning Questionnaire (Technical Report 91-B-004). *The Regents of the University of Michigan*. Retrieved from https://files.eric.ed.gov/fulltext/ ED338122.pdf

Plomin, R., & Petrill, S. A. (1997). Genetics and intelligence: What's new? *Intelligence*, *24*(1), 53-77. Retrieved from https://doi.org/10.1016/S0160-2896(97)90013-1

Pollard, C., & Pollard, R. (2004). Research Priorities in Educational Technology, *Journal of Research on Technology in Education, 37*(2), 145-160. Retrieved from https://doi.org/10.1080/15391523.2004.10782430

Pritchard, D. (2018). *What is this thing called Knowledge?* (4th ed.). Routledge, London.

Procee, H. (2006). Reflection in education: A Kantian epistemology. *Educational Theory, 56*(3), 237-253. Retrieved from https://doi.org/10.1111/j.1741-5446.2006.00225.x

Qi, J., Buechley, L., Ng, P., Cross, S., & Paradiso, J. A. (2018, April). Chibitronics in the Wild: Engaging New Communities in Creating Technology with Paper Electronics. In Proceedings of the 2018 *CHI Conference on Human Factors in Computing Systems*, Paper 2, ACM.

Raadschelders, J. C. (2011). The future of the study of public administration: embedding research object and methodology in epistemology and ontology. *Public Administration Review, 71*(6), 916-924. Retrieved from https://doi.org/10.1111/j.1540-6210.2011.02433.x

Reber, P. J., Knowlton, B. J., & Squire, L. R. (1996). Dissociable properties of memory systems: differences in the flexibility of declarative and nondeclarative knowledge. *Behavioral neuroscience, 110*(5), 861.

Reeves, J. L., Gunter, G. A., & Lacey, C. (2017). Mobile Learning in Pre-Kindergarten: Using Student Feedback to Inform Practice. *Educational Technology & Society, 20*(1), 37-44.

Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex Differences in Mathematics and Science Achievement: A Meta-Analysis of National Assessment of Educational Progress Assessments. Retrieved from http://dx.doi.org/10.1037/edu0000012

Repovš, G., & Baddeley, A. (2006). The Multi-Component Model of Working Memory: Explorations in Experimental Cognitive Psychology. *Neuroscience, 139,* 5-21. Retrieved from https://doi.org/10.1016/j.neuroscience.2005.12.061

Rey, G. D., & Fischer, A. (2013). The expertise reversal effect concerning instructional explanations. *Instructional Science, 41*(2), 407-429. Retrieved from https://doi.org/10.1007/s11251-012-9237-2

Richardson, J. T., Richardson, J., Engle, R., Hasher, L., Logie, R., Stoltzfus, E., & Zacks, R. (1996). Evolving concepts of working memory. *Working memory and human cognition*, 3-30. New York, NY: Oxford University Press.

Riding, R., & Douglas, G. (1993). The effect of cognitive style and mode of presentation on learning performance. *British Journal of Educational Psychology, 63*(2), 297-307.

Robert, M., & Savoie, N. (2006). Are there gender differences in verbal and visuospatial working-memory resources? *European Journal of Cognitive Psychology, 18*(03), 378-397. Retrieved from https://doi.org/10.1080/09541440500234104

Roberts, N. (2016). *The school curriculum and SATs in England: reforms since 2010. Number 06798, 26 February, 2016*. Retrieved from http://dera.ioe.ac.uk/25606/8/SN06798_Redacted.pdf

Roodenrys, S., Hulme, C., & Brown, G. (1993). The Development of Short-Term Memory Span: Separable Effects of Speech Rate and Long-Term Memory. *Journal of Experimental Child Psychology*, *56*(3), 431-442.

Rosenthal, R., & Rosnow, R.L. (2009). *Artifacts in Behavioral Research.* New York, NY: Oxford University Press.

Rovai, A. P., Baker, J. D., & Ponton, M. K. (2014). *Social Science Research Design and Statistics. A practitioner's guide to research methods and IBM SPSS analysis.* Chesapeake, VA: Watertree Press LLC.

Ryan, R. M., & Deci, E. L. (2000a). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. Retrieved from http://dx.doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2000b). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54-67. Retrieved from https://doi.org/10.1006/ceps.1999.1020

Sadoski, M., Goetz, E.T. & Fritz, J.B. (1993). Impact of Concreteness on Comprehensibility, Interest, and Memory for Text: Implications for Dual Coding Theory and Text Design. *Journal of Educational Psychology, 85*(2). Retrieved from http://psycnet.apa.org/fulltext/1993-32227-001.pdf

Saido, G.M., Siraj, S., Nordin, A.B.B., & Al Almedy, O. S. (2015). Higher Order Thinking Skills Among Secondary School Students in Science Learning. *MOJES: Malaysian Online Journal of Educational Sciences, 3*(3), 13-20. Retrieved from: http://mojes.um.edu. my/index.php/MOJES/article/view/12778

Saine, P. (2012). iPods, iPads, and the SMARTBoard: Transforming literacy instruction and student learning. *New England Reading Association Journal*, *47*(2), 74-81.

Savill, N., Ellis, A. W., & Jefferies, E. (2017). Newly-acquired words are more phonologically robust in verbal short-term memory when they have associated semantic representations. *Neuropsychologia*, *98*, 85-97. Retrieved from https://doi.org/10.1016/j.neuropsychologia.2016.03.006

Sayer, A. (1992). *Methods in Social Science, a Realist Approach* (2nd ed.). London: Routledge.

Schamber, L. (2000). Time-line interviews and inductive content analysis: their effectiveness for exploring cognitive behaviors. *Journal of the Association for Information Science and Technology, 51*(8), 734-744. Retrieved from https://doi.org/10.1002/(SICI)1097-4571(2000)51:8<734::AID-ASI60>3.0.CO;2-3

Schlag, S., & Ploetzner, R. (2011). Supporting learning from illustrated texts: conceptualizing and evaluating a learning strategy. *Instructional Science, 39*(6), 921-937. Retrieved from https://doi.org/10.1007/s11251-010-9160-3

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York, NY: Basic books.

Schweppe, J., & Rummer, R. (2014). Attention, Working Memory, and Long-Term Memory in Multimedia Learning: An Integrated Perspective Based on Process Models of Working Memory. *Educational Psychology Review*, *26*, 285-306. Retrieved from https://doi.org/10.1007/s10648-013-9242-2

Sclafani, S. (2002). No child left behind. *Issues in Science and Technology*, *19*(2), 43-47.

Scotland, J. (2012). Exploring the Philosophical Underpinnings of Research: Relating Ontology and Epistemology to the Methodology and Methods of the Scientific, Interpretive, and Critical Research Paradigms. *English Language Teaching*, *5*(9), 9-16. Retrieved from http://dx.doi.org/10.5539/elt.v5n9p9

Shaffer, D. R., & Kipp, K. (2013). *Developmental Psychology Childhood & Adolescence* (9th ed.). Canada: Jon-David Hague.

Sharples, M. (2000). The Design of Personal Mobile Technologies for Lifelong Learning. *Computers & Education*, *34*(3), 177-193. Retrieved from https://doi.org/10.1016/S0360-1315(99)00044-5

Sharples, M. (2013). Mobile Learning: Research, Practice and Challenges. *Distance Education in China*, *3*(5), 5–11. Retrieved from http://oro.open.ac.uk/37510/2/sharples.pdf

Sharples, M., Taylor, J., & Vavoula, G. (2005). Towards a theory of mobile learning. In *Proceedings of mLearn 2005 conference 25–28 October*. Cape Town, South Africa.

Shenk, D. (2017). What is the Flynn Effect, and How Does It Change Our Understanding of IQ? *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(1-2). Retrieved from https://doi.org/10.1002/wcs.1366

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, *2*(4), 419-436.

Schunk, D. H. (2012). *Learning theories an educational perspective* (6th ed.). Pearson.

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning, 2*(1). Available from http://jotamac.typepad.com/jotamacs_weblog/files/connectivism.pdf

Skinner, B.F. (1965). *Science and Human Behavior*. New York: The Macmillan Company. (Original work published 1953).

Skinner, B.F. (1954). The Science of Learning and the Art of Teaching. In E. Stones (Ed.), *Readings in Educational Psychology, Learning and Teaching* (pp. 301-312). Abingdon, Oxon: Routledge.

Smart, M.P. (2008). The word and the world: technology aids English language learners. Edutopia. Retrieved on 11 February, 2018, from http://www.edutopia.org/technology-software-english-languagelearners

Smuts, A. (2009). What Is Interactivity? *The Journal of Aesthetic Education*, *43*(4), 53-73. Retrieved from http://www.jstor.org/stable/25656247

So, K., & Kang, J. (2014). Curriculum reform in Korea: Issues and challenges for twenty-first century learning. *The Asia-Pacific Education Researcher*, *23*(4), 795-803. Retrieved from https://doi.org/10.1007/s40299-013-0161-2

Spear, N. E., & Riccio, D. C. (1994). *Memory: Phenomena and principles.* US: Allyn & Bacon.

Spearman, C. (1927). *The abilities of man*. New York, NY: Cambridge.

Speed, B. (1991). Reality exists OK? An argument against constructivism and social constructionism. *Journal of Family Therapy, 13*(4), 395-409. Retrieved from https://doi.org/10.1046/j..1991.00436.x

Spiegel, A., & Rodríguez, G. (2016). Students at University have Mobile Technologies. Do they do m-learning? *Procedia-Social and Behavioral Sciences*, *217*, 846-850. Retrieved from https://doi.org/10.1016/j.sbspro.2016.02.006

Squire, L. R. & Zola-Morgan, S. (1988). Memory: brain systems and behaviour. *Trends in Neurosciences, 11*(4), 170-175. Retrieved from https://doi.org/10.1016/0166-2236(88)90144-0

St Clair-Thompson, H. L., & Gathercole, S.E. (2006). Executive Functions and Achievements in School: Shifting, Updating, Inhibition, and Working Memory. *Quarterly Journal of Experimental Psychology*, *59*(4), 745-759. Retrieved from http://journals.sagepub.com/doi/pdf/10.1080/17470210500162854

St Clair-Thompson, H. L., & Holmes, J. (2008). Improving short-term and working memory: Methods of memory training. *New research on short-term memory,* 125-154. Retrieved from http://lucid-research.net/documents/research/BookChapter_Acad_NewResearchOnShortTermMemory08_ImprovingMethodsOfMemoryTraining.pdf

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York, NY: Cambridge University Press.

Stickgold, R., & Walker, M. P. (2007). Sleep-dependent memory consolidation and reconsolidation. *Sleep Medicine*, *8*, 331-343. Retrieved from https://doi.org/10.1016/j.sleep. 2007.03.011

Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within-and across-nation assessment of 10 years of PISA data. *PloS one*, *8*(3), e57988. Retrieved from https://doi.org/10.1371/journal.pone.0057988

Stroebe, W., Gadenne, V., & Nijstad, B. A. (2018). Do Our Psychological Laws Apply Only to College Students?: External Validity Revisited. *Basic and Applied Social Psychology,* 1-12.

Strike, K. A. (2012). The Ethics of Educational Research. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of Complementary Methods in Educational Research* (pp. 57-74). USA: Lawrence Erlbaum Associates Publishers.

Subrahmanyam, S. (2017). *Retain: building a concept recommendation system that leverages spaced repetition to improve retention in educational settings*. Thesis from the University of Illinois. Retrieved from http://hdl.handle.net/2142/ 97490

Svendsen, D. (1983). Factors Related to Changes in IQ: A Follow-Up Study of Former Slow Learners. *Journal of Child Psychology and Psychiatry, 24*(3), 405-413. Retrieved from https://doi.org/10.1111/j.1469-7610.1983.tb00117.x

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science, 12*(2), 257-285. Retrieved from https://doi.org/10.1207/s15516709cog 1202_4

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction, 4*(4), 295-312. Retrieved from https://doi.org/10.1016/ 0959-4752(94)90003-5

Sweller, J. (2008). Cognitive Load Theory and the Use of Educational Technology. *Educational Technology*, *48*(1), 32-35.

Sweller, J. (2010). Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review*, *22*(2), 123-138.

Sweller, J., Van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, *10*(3), 251-296. Retrieved from https://doi.org/10.1023/A:1022193728205

Tabatabaei, S., Ahadi, H., Bani-Jamali, S., Bahrami, H., & Khamesan, A. (2017). The Effects of Motivated Strategies for Learning Questionnaire (MSLQ) on Students' Cognitive and Meta-Cognitive Skills. *NeuroQuantology, 15*(2), 239-245. Retrieved from doi: 10.14704/nq.2017.15.2.1068

Takooshian, H., Gielen, U. P., Plous, S., Rich, G. J., & Velayo, R. S. (2016). Internationalizing undergraduate psychology education: Trends, techniques, and technologies. *American Psychologist, 71*(2), 136-147.

Terry, W. S. (2016). *Learning and Memory: Basic Principles, Processes, and Procedures* (4th ed.). Abingdon, Oxon: Routledge.

Texas Instruments. (2018, 14 August). Understanding flap actuators and what drives them in automotive HVAC systems. Retrieved from https://e2e.ti.com/blogs_/b/behind_the_wheel/archive/2017/04/13/understanding-flap-actuators-and-what-drives-them-in-automotive-hvac-systems

Tham, J. C. K. (2018). Interactivity in an age of immersive media: Seven dimensions for wearable technology, internet of things, and technical communication. *Technical Communication*, *65*(1), 46-65.

The National Curriculum (2018). *The Government UK*. Retrieved from https://www.gov.uk/national-curriculum

Thomas, G. (2017). *How to Do Your Research Project: A Guide for Students*. Sage.

Tobin, K. (2009). Repetition, Difference, and Rising Up with Research in Educational. In K. Ercikan, & W. M. Roth (Eds.), *Generalising in Educational Research, Beyond Qualitative and Quantitative Polarization* (pp. 149-160). New York, NY: Routledge.

Tomaszewski, B., Vodacek, A., Parody, R., & Holt, N. (2014). Spatial Thinking Ability Assessment in Rwandan Secondary Schools: Baseline Results. *Journal of Geography, 114*(2), 39-48. Retrieved from https://doi.org/10.1080/00221341.2014.918165

Tondeur, J., van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: a systematic review of qualitative evidence. *Educational Technology Research and Development*, *65*(3), 555-575. Retrieved from https://link.springer.com/content/pdf/10.1007%2Fs11423-016-9481-2.pdf

Tourva, A., Spanoudis, G., Demetriou, A. (2016). Cognitive correlates of developing intelligence: The contribution of working memory, processing speed and attention. *Intelligence*, *54*, 136-146. Retrieved from https://doi.org/10.1016/j.intell.2015.12.001

Traxler, J. (2011). Research Essay: Mobile Learning. *International Journal of Mobile and Blended Learning (IJMBL), 3*(2), 57-67.

Tulving, E. (1972). Episodic and semantic memory. *Organization of Memory*, 1, 381-403.

UK Government (2018). *Childcare and Parenting, School and Education, National Curriculum, Key Stage 3 and 4.* Retrieved from https://www.gov.uk/national-curriculum/key-stage-3-and-4

Van Der Elst, W. I. M., Van Boxtel, M. P., Van Breukelen, G. J., & Jolles, J. (2005). Rey's verbal learning test: normative data for 1855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society*, *11*(3), 290-302.

van der Meij, H., & van der Meij, J. (2014). A comparison of paper-based and video tutorials for software learning. *Computers & education*, *78*, 150-159. Retrieved from https://doi.org/10.1016/j.compedu.2014.06.003

Vandewaetere, M., & Clarebout, G. (2013). Cognitive load of learner control: extraneous or germane load? *Education Research International*, 2013. Retrieved from http://dx. doi.org/10.1155/2013/902809

Van Merriënboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational psychology review*, *17*(2), 147-177. doi:10.1007/s10648-005-3951-0

Van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let’s learn from cows in the rain. *PloS one, 8*(7), e68967.

Veenman, M. V., Van Hout-Wolters, B. H., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and learning*, *1*(1), 3-14. doi:10.1007/s11409-006-6893-0

Vellinga, A., Cormican, M., Hanahoe, B., Bennett, K., & Murphy, A. W. (2011). Opt-out as an acceptable method of obtaining consent in medical research: a short report. *BMC medical research methodology*, *11*(1), 40. doi:10.1186/1471-2288-11-40

Volante, L. (2015). The Impact of PISA on Education Governance: Some Insights from Highly Reactive Policy Contexts. *International Studies in Educational Administration (Commonwealth Council for Educational Administration & Management (CCEAM))*, *43*(2), 103-117.

Von Glasersfeld, E. (1989). *Knowing without Metaphysics: Aspects of the Radical Constructivist Position*.

Von Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Learning. Studies in Mathematics Education Series: 6.* Taylor & Francis Inc: Bristol.

Vygotsky, L. S. (1978). *Mind in Society*. Harvard University Press, Cambridge.

Walliman, N. & Buckler, S. (2008). *Your Dissertation in Education*. London: SAGE Publications Ltd.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational technology research and development*, *53*(4), 5-23. Retrieved from https://link.springer.com/content/pdf/10.1007%2FBF02504682. Pdf

Wang, C. & Liou, P. (2017). Students' motivational beliefs in science learning, school motivational contexts, and science achievement in Taiwan. *International Journal of Science Education, 39*(7), 898-917. doi: 10.1080/09500693.2017.1310410

Wang, Y. S., Wu, M. C., & Wang, H. Y. (2009). Investigating the determinants and age and gender differences in the acceptance of mobile learning. *British journal of educational technology, 40*(1), 92-118. Retrieved from https://doi.org/10.1111/j. 1467-8535.2007.00809.x

Wästlund, E., Reinikka, H., Norlander, T., & Archer, T. (2005). Effects of VDT and paper presentation on consumption and production of information: Psychological and physiological factors. *Computers in Human Behavior*, *21*(2), 377-394. doi:10.1016/j.chb.2004.02.007

Webster, D. (2017). *Prisoners of Geography: Ten Maps That Explain Everything About the World.* Elliott & Thompson

Weimer, M. (2002). *Learner-centered teaching: Five key changes to practice*. San Francisco, CA: Jossey-Bass.

Werning, M. & Cheng, S. (2017). Taxonomy and Unitary of Memory. In S. Bernecker & K. Michaelian (Eds.), *The Routledge Handbook of Philosophy of Memory* (pp. 7-20). Abingdon: Routledge.

White, J. (2008). Illusory intelligences? *Journal of Philosophy of Education*, *42*(3-4), 611-630. Retrieved from https://doi.org/10.1111/j.1467-9752.2008.00643.x

Wiburg, K. M. (1995). An Historical Perspective on Instructional Design: Is it Time to Exchange Skinner's Teaching Machine for Dewey's Toolbox? *Proceeding CSCL '95 The first international conference on Computer support for collaborative learning,* 385-391. Retrieved from http://michel.delord.free.fr/bibli1998/instructional_design.html

Wilhelm, I., Prehn-Kristensen, A., & Born, J. (2012). Sleep-dependent memory consolidation–what can be learnt from children? *Neuroscience & Biobehavioral Reviews, 36*(7), 1718-1728. doi:10.1016/j.neubiorev.2012.03.002

Wilson, D. I., & Maclaren, P. (2013). From chalk talk to tablet talk: pedagogies for control engineering. *IFAC Proceedings Volumes, 46*(17), 144-149. Retrieved from https://doi.org/10.3182/20130828-3-UK-2039.00036

Wilson, E. (2017). *School-based research: A guide for education students* (3rd ed.). London: Sage.

Wingkvist, A., & Ericsson, M. (2013). A Survey of Research Methods and Purposes in Mobile Learning. In D. Parsons (Ed.). *Innovations in Mobile Educational Technology and Applications* (pp. 14-31). USA: Information Science Reference.

Wolters, C. A., & Pintrich, P. R. (1998). Contextual differences in student motivation and self-regulated learning in mathematics, English, and social studies classrooms. *Instructional Science*, *2*6(1-2), 27-47. Retrieved from https://doi.org/10.1023/A:1003035929216

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction, 22*(6), 449-457. doi:10.1016/j.learninstruc.2012.05.004

Wong, A., Marcus, N., Ayres, P., Smith, L., Cooper, G. A., Paas, F., & Sweller, J. (2009). Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior*, *25*(2), 339-347. doi:10.1016/j.chb.2008.12.012

World Health Organisation (2015). *Core Competencies in Adolescent Health and Development for Primary Care Providers Including a Tool to Assess the Adolescent Health and Development Component in Pre-Service Education of Health-Care Providers.* Geneva: WHO Press.

Wynn, D., & Williams, C. K. (2012). Principles for Conducting Critical Realist Case Study Research in Information Systems. *MIS Quarterly 36*(3), 787-810.

Yeh, Y. C. (2009). Integrating e-learning into the Direct-instruction Model to enhance the effectiveness of critical-thinking instruction. *Instructional Science*, *37*(2), 185-203. doi:10.1007/s11251-007-9048-z

Yeoman, K., Bowater, L., & Nardi, E. (2016). The representation of scientific research in the national curriculum and secondary school pupils' perceptions of research, its function, usefulness and value to their lives [version 2; referees: 2 approved]. *F1000Research, 4*:1442. doi: 10.12688/f1000research.7449.2

Yim, M. Y. C., Chu, S. C., & Sauer, P. L. (2017). Is augmented reality technology an effective tool for e-commerce? An interactivity and vividness perspective. *Journal of Interactive Marketing*, *39*, 89-103. Retrieved from https://www.sciencedirect.com/science/article/pii/S1094996817300336

Yung, H. I., & Paas, F. (2015a). Effects of Computer-Based Visual Representation on Mathematics Learning and Cognitive Load. Journal of *Educational Technology & Society*, *18*(4), 70-77.

Yung, H. I., & Paas, F. (2015b). Effects of Cueing by a Pedagogical Agent in an Instructional Animation: A Cognitive Load Approach. *Educational Technology & Society*, *18*(3), 153-160.

Yusuf, M. O. (2005). Information and Communication Technology and Education: Analysing the Nigerian National Policy for Information Technology. *International Education Journal*, *6*(3), 316-321. Retrieved from https://files.eric.ed.gov/fulltext/EJ854985.pdf

Zachariadis, M., Scott, S. V., & Barrett, M. I. (2013). Methodological Implications of Critical Realism for Mixed-Methods Research. *MIS quarterly*, *37*(3), 855-879. Retrieved from https://pdfs.semanticscholar.org/5b97/6a91750cc2a2a53e616a09fc2042f50c00e6.pdf

Ziefle, M. (1998). Effects of display resolution on visual performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *40*(4), 554-568. Retrieved from http://journals.sagepub.com/doi/pdf/10.1518/001872098779649355

Appendix A – Ethical Approval and School Documents

University of Reading, Institute of Education

Ethical Approval Form A (version September 2013)

Tick one:

Staff project: ___ PhD: ____ **EdD**: ✓

Name of applicant (s): Ms Denise Anne Thompson (nee Erlank)

Title of project: To examine the process of learning from a multidisciplinary perspective, using extended multimedia tasks on technology and paper-based exercises, to identify the prime cognitive load for successful encoding and retrieval from long-term memory.

Name of supervisors: Dr Daisy Powell, Professor Rhona Stainthorp

Please complete the form below including relevant sections overleaf.

	YES	NO
Have you prepared an Information Sheet for participants and/or their parents/carers that:		
a) explains the purpose(s) of the project	✓	
b) explains how they have been selected as potential participants	✓	
c) gives a full, fair and clear account of what will be asked of them and how the information that they provide will be used	✓	
d) makes clear that participation in the project is voluntary	✓	
e) explains the arrangements to allow participants to withdraw at any stage if they wish	✓	
f) explains the arrangements to ensure the confidentiality of any material collected during the project, including secure arrangements for its storage, retention and disposal	✓	
g) explains the arrangements for publishing the research results and, if confidentiality might be affected, for obtaining written consent for this	✓	
h) explains the arrangements for providing participants with the research results if they wish to have them	✓	
i) gives the name and designation of the member of staff with responsibility for the project together with contact details, including email . If any of the project investigators are students at the IoE, then this information must be included and their name provided	✓	

k) explains, where applicable, the arrangements for expenses and other payments to be made to the participants	N/A		
j) Includes a standard statement indicating the process of ethical review at the University undergone by the project, as follows: 'This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct'.	✓		
k) includes a standard statement regarding insurance: "The University has the appropriate insurances in place. Full details are available on request".	✓		
Please answer the following questions			
1) Will you provide participants involved in your research with all the information necessary to ensure that they are fully informed and not in any way deceived or misled as to the purpose(s) and nature of the research? (Please use the subheadings used in the example information sheets on blackboard to ensure this).	✓		
2) Will you seek written or other formal consent from all participants, if they are able to provide it, in addition to (1)?	✓		
3) Is there any risk that participants may experience physical or psychological distress in taking part in your research?		✓	
4) Have you taken the online training modules in data protection and information security (which can be found here: http://www.reading.ac.uk/internal/imps/Staffpages/imps-training.aspx)?	✓		
5) Have you read the Health and Safety booklet (available on Blackboard) and completed a Risk Assessment Form to be included with this ethics application?	✓		
6) Does your research comply with the University's Code of Good Practice in Research?	✓		
	YES	NO	N.A.
7) If your research is taking place in a school, have you prepared an information sheet and consent form to gain the permission in writing of the head teacher or other relevant supervisory professional?	✓		
8) Has the data collector obtained satisfactory DBS clearance?	✓		
9) If your research involves working with children under the age of 16 (or those whose special educational needs mean they are unable to give informed consent), have you prepared an information sheet and consent form for parents/carers to seek permission in writing, or to give parents/carers the opportunity to decline consent?	✓		
10) If your research involves processing sensitive personal data, or if it involves audio/video recordings, have you obtained the explicit consent of participants/parents?			✓
11) If you are using a data processor to subcontract any part of your research, have you got a written contract with that contractor which (a) specifies that the contractor is required to act only on your instructions, and (b) provides for appropriate technical and organisational security measures to protect the data?			✓
12a) Does your research involve data collection outside the UK? Yes, but the researcher is resident within that country and is insured by the employing institution.	✓		
12b) If the answer to question 11a is "yes", does your research comply with the legal and ethical requirements for doing research in that country?			N/A
13a. Does the proposed research involve children under the age of 5?		✓	

13b. If the answer to question 12a is "yes": My Head of School (or authorised Head of Department) has given details of the proposed research to the University's insurance officer, and the research will not proceed until I have confirmation that insurance cover is in place: the researcher is resident within that country and is insured by the employing institution.			N/A
If you have answered YES to Question 3, please complete Section B below			

PLEASE COMPLETE **EITHER SECTION A OR B** AND PROVIDE THE DETAILS REQUIRED IN SUPPORT OF YOUR APPLICATION, THEN SIGN THE FORM (SECTION C)

A: My research goes beyond the 'accepted custom and practice of teaching' but I consider that this project has **no** significant ethical implications.	✓

Give a brief description of the aims and the methods (participants, instruments and procedures) of the project in up to 200 words. Attach any consent form, information sheet and research instruments to be used in the project (e.g. tests, questionnaires, interview schedules).

Please state how many participants will be involved in the project:
Approximately **430 participants**

This form and any attachments should now be submitted to the Institute's Ethics Committee for consideration. Any missing information will result in the form being returned to you.

Aim: To investigate the process of learning to ascertain the following: whether cognitive load affects learning (Mayer, 2008); whether recall is greater when assessed using the same resource on which learning took place (Tulving & Thomson, 1973); and whether individual differences play a part in recall (Skinner, 1965). The research will examine Mayer's Multimedia Principle, which states that learning is improved with combined words and pictures than from words alone; and the Redundancy Principle, which states that learning is improved with reduced input as the sensory channel can be overloaded, based on Moreno's Cognitive-Affective theory of Learning with Media (Mayer, 2009; Moreno, 2006). All KS3 students in a British curriculum international school in Doha will complete a short, paper-based learning activity in the classroom, which will be used as the control against which the success of learning, using different combinations of multimedia technology (i.e. graphics, sound, animation/video, and text) will be measured. The research will hopefully build on the understanding of, and help to validate, existing theories of memory and learning.

Methods: Participants**:** 430 Year 7-9 students, male and female.

Instruments:

There will be an initial assessment of students using the Automated Working Memory Assessment (AWMA) (http://www.pearsonclinical.co.uk/Psychology/ChildCognitionNeuropsychologyandLanguage/ChildMemory/AutomatedWorkingMemoryAssessment(AWMA)/AutomatedWorkingMemoryAssessment(AWMA).aspx) and an adapted Motivated Strategies for Learning Questionnaire (Pintrich, & DeGroot., 1990). Students' data which includes CAT scores and reading ages (Access Reading Test) will also be obtained from the school with parents' consent.

Parents and students will be informed that all students will take part in the research activities, and if they do not give their consent their data will be omitted from the research. Parental permission will be gained in the letter to parents and students' consent will be gained in the initial questionnaire.

Each condition requires a short prior knowledge assessment, after which the learning condition will take place. The learning conditions comprise of: a one page paper exercise on a topic (e.g. brain); a tablet using Apps/documents that will have different combinations of graphics, text, audio, and animation for topics 2 (e.g. heart) and 3 (e.g. map directions). The prior knowledge assessment will be given again at the end of each learning condition to assess short-term memory retention.

An initial Questionnaire will be given to participants within a week of the final learning condition to gauge their impressions of the research. There will also be a repeat of the prior knowledge assessments, in one session, to determine long-term memory retention six weeks after the final learning condition, when the End of Research Questionnaire will be given to assess any further studying of the topics which could impact on the results for long-term memory recall.

Procedures: natural environment–the classroom.

Two tutor groups from each year group (7-9) will be randomly assigned to each of three conditions. The first condition for all students will be paper-based; conditions 2 and 3 will use technology (iPad) using different combinations of graphics, text, audio and animation for different groups. Participants will be given a prior knowledge assessment lasting 5 minutes. They will then be given the testing material which they will work with for 5 minutes after which a normal lesson i.e. ICT will follow. Participants will complete the assessment form again at the end of the hour lesson to assess short-term memory retention. Participants will be given the same form in May of the summer term to determine long-term memory retention. They will also complete a questionnaire to assess their thoughts on the process and to determine whether they studied the content of the topics between the short-term and long-term memory assessments.

Students/teachers/parents will not be told that the research is examining long-term memory; however, they have been properly informed of the kind of activities students will be asked to participate in.

B: I consider that this project **may** have ethical implications that should be brought before the Institute's Ethics Committee.

Please provide all the further information listed below in a separate attachment.

1. title of project
2. purpose of project and its academic rationale
3. brief description of methods and measurements
4. participants: recruitment methods, number, age, gender, exclusion/inclusion criteria
5. consent and participant information arrangements, debriefing (attach forms where necessary)
6. a clear and concise statement of the ethical considerations raised by the project and how you intend to deal with then.
7. estimated start date and duration of project

This form and any attachments should now be submitted to the Institute's Ethics Committee for consideration. Any missing information will result in the form being returned to you.

C: SIGNATURE OF APPLICANT:

I have declared all relevant information regarding my proposed project and confirm that ethical good practice will be followed within the project.

Signed Print Name: Denise Thompson

Date: 27 October, 2014

STATEMENT OF ETHICAL APPROVAL FOR PROPOSALS SUBMITTED TO THE INSTITUTE ETHICS COMMITTEE

This project has been considered using agreed Institute procedures and is now approved.

Signed: Print Name A J Kempe

Date 2.12.14

(IoE Research Ethics Committee representative)*

* A decision to allow a project to proceed is not an expert assessment of its content or of the possible risks involved in the investigation, nor does it detract in any way from the ultimate responsibility which students/investigators must themselves have for these matters. Approval is granted on the basis of the information declared by the applicant.

University of Reading

Parent/carer information sheet

Research Project: To investigate optimal conditions for learning.

Research Team: Researcher: Ms Denise Thompson
Supervisors: Dr Daisy Powell
Professor Rhona Stainthorp,

Dear Parents

We would like to inform you of a research study that will be taking place within the school about learning using technology (iPads). The study is being conducted as a research project towards an Educational Doctorate.

What is the study?
The study will investigate the optimal conditions for learning using different combinations of media i.e. graphics, text, audio, and animation through the use of technology (iPads). It hopes to make recommendations regarding how we can best help students to make progress in learning.

Why has my child been chosen to take part?
Your child has been chosen to take part in the project because he/she is in Key Stage 3 (Year 7-9). All learners in KS3 have been selected to take part.

Does my child have to take part?
Your child would need to take part in the activities, which is in line with what they do in regular lessons. However, you may choose to have your child's data excluded from the research at any time during the project, without any repercussions to you or your child, by contacting the researcher: dethompson@xxx.com

What will happen when my child takes part?
A Working Memory assessment (15 minutes) and Motivated Strategies for Learning Questionnaire (20 minutes) will be completed by students initially.

Students will complete a short textbook style exercise to determine whether technology has an effect on learning; additional short exercises will use a combination of media on the iPad. Sessions, which will take place on different days, should take about 20 minutes each, of which there should be a total of four. At the end of the research students will need to complete a questionnaire to give their impressions of the study. All tasks will take place during the normal school day.

This study may lead on to a further longitudinal study to identify the effects of using technology in learning, or to gain parents' perspectives. In this case we would contact you again to gain further consent. Finally, we would like your permission for School xx to pass on details of your child's information which includes their CAT score and reading age; this will be held in the strictest confidence.

What are the risks and benefits of taking part?
The information you and your child give will remain confidential and will only be seen by the research team listed at the start of this letter. Neither you, your child, nor the school will be identifiable in any published report resulting from the study. Taking part will in no way influence the grades your child receives at school. Information about individuals, except working memory, will not be shared with the school.

Participants in similar studies have found it interesting to take part. We anticipate that the findings of the study will be useful for teachers using technology in lessons.

What will happen to the data?
Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. No identifiers linking you, your child or the school to the study will be included in any sort of report that might be published. Children will be assigned a number and will be referred to by that number. Research records will be stored securely and on a password-protected computer: only the researcher and supervisors will have access to the records. The data will be destroyed securely once the findings of the study are written up, after five years. The results of the study may be presented at national and international conferences, and in written reports and articles.

Who has reviewed the study?
This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct.

What happens if I/ my child change our mind?
You can change your mind at any time without any repercussions and your child's data will be omitted.

What happens if something goes wrong?
In the unlikely case of concern or complaint, you can contact Dr Daisy Powell, University of Reading; Tel: , email:

Where can I get more information?
If you would like more information, please contact Ms Denise Thompson at xxx.
Tel: email:

We do hope that you will agree to your child's data being used. We will assume that you give your permission unless we hear from you. Only if you **do not** give your permission, please complete the attached form and return it to your child's tutor before the end of the autumn term i.e. Wednesday, 17th December. Data collection will take place during the spring term.

Thank you for your time.
Yours sincerely,

Denise Thompson
ICT and Computer Science teacher
EdD student, University of Reading

Research Project: To investigate optimal conditions for learning.

DO NOT RETURN THIS FORM IF YOU GIVE CONSENT FOR YOUR CHILD'S DATA TO BE USED

Parent/Carer Withdrawal Form

I have received a copy of it the Information Sheet about the project and have read the contents and DO NOT WANT my child's data to be included in the research.

Name of Child: ____ **Tutor Group**: _____

Name of school: xx

Please tick to confirm withdrawal of your child's data from the research:

I do not consent to my child's data being included in the research project. ☐

Name of parent/carer: ____

Signed:____________________________________ **Date**: _____________________

Please return to your child's tutor who will pass it to Ms Denise Thompson.

Thank you.

Principal and Head Teacher information sheet

University of Reading

Research Project: To investigate optimal conditions for learning.

Research Team: Researcher: Ms Denise Thompson
Supervisors: Dr Daisy Powell
Professor Rhona Stainthorp,

Dear XX and XX

I am writing to invite your school to take part in a research study about the use of technology in learning.

What is the study?

A few weeks ago I spoke with you about a study I would like to conduct at School xx to collect data for my research thesis for the University of Reading, the study being conducted as a research project towards an Educational Doctorate. The research aims to investigate two competing theories about learning: Mayer's Multimedia Principle which states learning is improved with combined words and pictures, and Redundancy Principle which states learning is improved with reduced input. As such, it will investigate the optimal conditions for learning using different combination of media in technology i.e. graphics, text, audio, and animation. It hopes to make recommendations regarding how we can best help students to make progress in learning.

Why has this school been chosen to take part?

This school has been selected as it uses both traditional methods of learning i.e. textbooks as well as technology i.e. iPads, as a matter of course.

Does the school have to take part?

It is entirely up to you whether you give permission for the school to participate. You may also withdraw your consent to participation at any time during the project, without any repercussions to you, by contacting the researcher or supervisors, emails above.

What will happen if the school takes part?

With your agreement, participation would involve Key Stage 3 students completing a Working Memory assessment (15 minutes) and a Motivated Strategies for Learning Questionnaire (20 minutes). These could be administered during an English lesson in the first instance, and an ICT lesson and tutor time in the second and third instances to reduce the amount of time missed in any one subject. Students would then be randomly assigned into groups by tutor group and undergo learning in different conditions – all would be assessed with a textbook style exercise, then with a combination of media using technology. In total this would take a maximum of one hour spread over three lessons. In addition, in the summer term I would like students to complete a final assessment to determine long-term memory retention and a questionnaire to gain student impressions of the study. This should take a total of 40 minutes which could be conducted in an ICT lesson (20 minutes) and tutor time (20 minutes).

Additionally, it would be helpful to have access to student CAT scores and reading ages. Assessment would be administered during the school day, overseen by Ms Denise

Thompson and any additional staff administering assessment will be given careful guidance on how to conduct the data collection.

This study may lead on to further study to identify the longitudinal effects of using technology in learning, or to gain parents' perspective.

If you agree to the school's participation, I will seek further consent from parents/carers and the children themselves, as well as from the subject leaders giving up class time to participate in the study.

What are the risks and benefits of taking part?

The information given by participants in the study will remain confidential and will only be seen by the researcher and supervisors listed at the start of this letter. Neither you, the children or the school will be identifiable in any published report resulting from the study. Information about individuals will not be shared with the school.

Participants in similar studies have found it interesting to take part. We anticipate that the findings of the study will be useful for teachers using technology in lessons.

What will happen to the data?

Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. The records of this study will be kept private. No identifiers linking you, the children or the school to the study will be included in any sort of report that might be published. Participants will be assigned a number and will be referred to by that number in all records. Research records will be stored securely in a locked filing cabinet and on a password-protected computer and only the research team will have access to the records. The data will be destroyed securely once the findings of the study are written up, after five years. The results of the study may be presented at national and international conferences, and in written reports and articles. We can send you electronic copies of these publications if you wish.

What happens if I change my mind?

You can change your mind at any time without any repercussions. If you change your mind after data collection has ended, we will discard the school's data.

What happens if something goes wrong?

In the unlikely case of concern or complaint, you can contact Dr Daisy Powell, University of Reading; Tel: email:

Where can I get more information?

If you would like more information, please contact either the researcher, D Thompson, or Dr Daisy Powell. Thank you for agreeing to take part in the study. To make it official, please complete the enclosed consent form and return it to Ms D Thompson.

This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct. The University has the appropriate insurances in place. Full details are available on request.

Thank you for your time.
Yours sincerely,
Denise Thompson

Head Teacher Consent Form – School xx

- ☐ I have read the Information Sheet about the project and received a copy of it.

- ☐ I understand what the purpose of the project is and what is required of me. All my questions have been answered.

Name of Principal: xx

Name of school: xxx

Please tick as appropriate:

I consent to the involvement of my school in the project as outlined in the Information Sheet ☐

Signed: ______________________________

Date: ______________________________

Student Pamphlet

What happens next?

Your parents have been sent a letter asking for their permission for you to take part in this project.

We will check with you that you are happy to help us with our project.

If you have any questions please contact

Ms Denise Thompson

IT3

dethompson@

Researcher

Ms Denise Thompson

University Supervisors

Dr Daisy Powell

Professor Rhona Stainthorp

This project has been reviewed following the procedures of the University of Reading Research Ethics Committee and has been given a favourable ethical opinion for conduct.

University of Reading

Institute of Education
London Road Campus
RG1 5EX

Research Project:

Finding the Best Way to Learn Better

Information Sheet

We are doing a project to understand what the best ways are of 'learning to remember'. We would like you to help us with the project. We have already asked your parents, Principal and Head Teacher if they are happy for you to help us.

Why have I been invited to take part?

You have been invited to take part because you are in Year 7, 8 and 9, because you use iPads in school, and because the School is interested in knowing how we can help you learn better.

What will I have to do if I agree to take part?

1. *We will assess:*
   - *Your working memory, a 15 minute exercise, and*
   - *Motivation, a 20 minute questionnaire.*
2. *We will give you an assessment (5 minutes) to see how much you already know about the activity topic.*
3. *Then, we will give you a short task to do from paper or iPad (5 minutes), and later assess how much you remember (5 minutes). We will do this four times.*
4. *Later on we will give you a questionnaire to see what you thought of it all.*

These will all take place during the school day.

Thank you very much to the ICT department, tutors, and the School for helping.

Will it affect my school levels?

Taking part is separate to your lessons and will not change your grades/levels at all.

Will anyone know about my answers?

- Your answers will be anonymous; no one will be able to identify you. Each student in Year 7, 8, and 9 will be given a number and results will be put next to that number.
- We will tell the school, you and your parents general results (how everyone has done), not individual results.

Will it help me if I take part?

We think you will find it interesting and fun to do the tasks. Your answers will help your teachers to know the best ways to teach you.

Do I have to take part?

You will need to do the tasks as part of your normal school work, but if you wish, we won't include your results in our research project. Also, you can stop helping us with our project at any time, without giving a reason. Just email me or ask your tutor or parents to email me if you want to stop.

Thank **you** for helping!

Appendix B – Heart Text-Only Condition

The Heart–Text Only (Paper, M-technology)

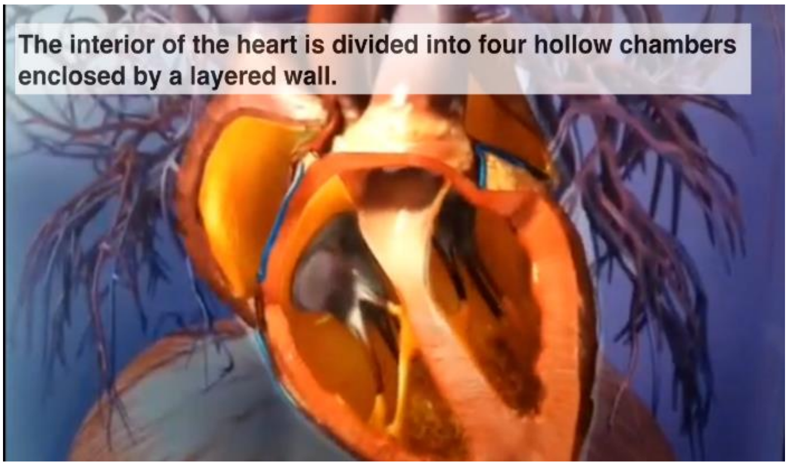
The Heart

1. The **heart** is a muscular **organ** that functions as a double **pump** to control **blood flow**.
2. The heart of an adult **male** is slightly **larger** than that of an adult female.
3. On average, a normal adult heart is about the **size of a fist** and weighs approximately **300g**; less than one pound.
4. The exterior of the heart has a **conical shape**; the interior of the heart is divided into **four hollow chambers** enclosed by a layered **wall**.
5. The heart has a **right** and a **left** side, each side with two chambers, an **atrium** and a **ventricle**. The right and left sides of the heart work in co-ordination **to pump oxygenated blood throughout the body**, and **de-oxygenated blood to the lungs**.
6. The heart of a healthy adult **beats** about **60-70 times a minute** to keep blood constantly moving. Within the thoracic cavity there is a thin sac known as the **pericardium** that encloses and protects the heart.

Appendix C – Heart Text, Graphics & Audio Condition

The Heart–Text, Graphics and Audio (M-technology)

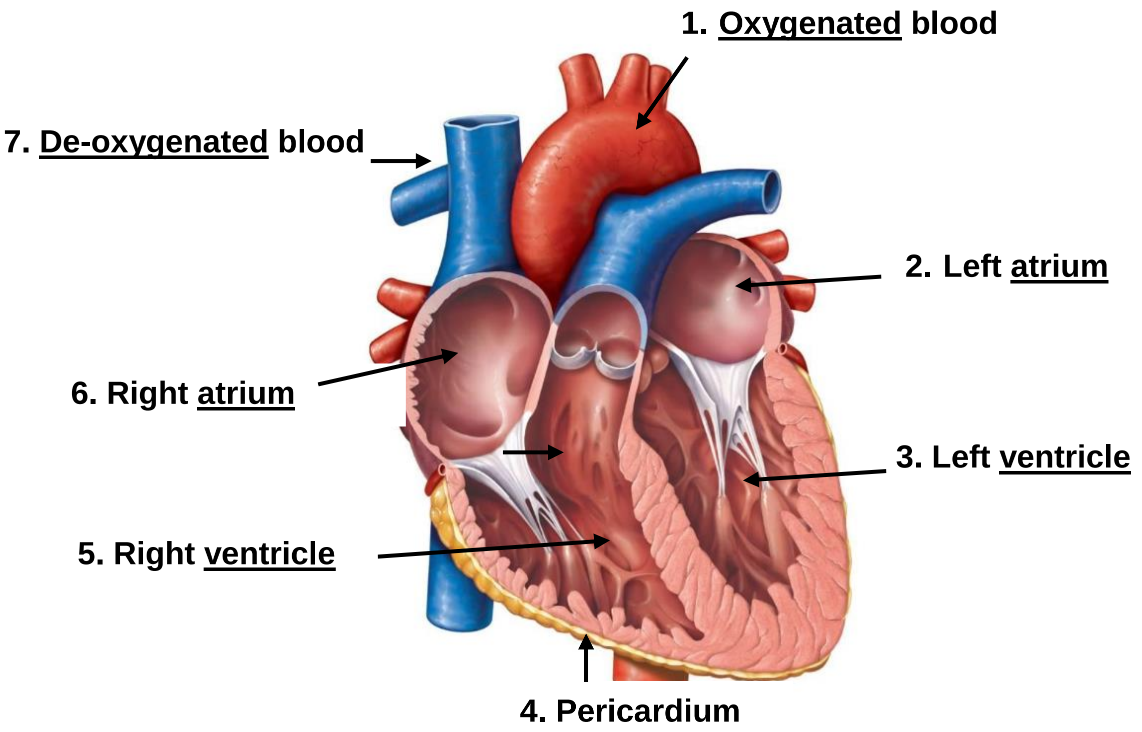
The graphics and audio condition did not have the text on the screen.



Appendix D – Heart Text & Graphics Condition

The Heart–Text and Graphics (Paper, M-technology)

The Heart



1. The **heart** is a muscular **organ** that functions as a double **pump** to control **blood flow**.
2. The heart of an adult **male** is slightly **larger** than that of an adult female.
3. On average, a normal adult heart is about the **size of a fist** and weighs approximately **300g**; less than one pound.
4. The exterior of the heart has a **conical shape**; the interior of the heart is divided into **four hollow chambers** enclosed by a layered **wall**.
5. The heart has a **right** and a **left** side, each side with two chambers, an **atrium** and a **ventricle**. The right and left sides of the heart work in co-ordination **to pump oxygenated blood throughout the body**, and **de-oxygenated blood to the lungs**.
6. The heart of a healthy adult **beats** about **60-70 times a minute** to keep blood constantly moving. Within the thoracic cavity there is a thin sac known as the **pericardium** that encloses and protects the heart.

Appendix E – Heart Assessment

The Heart Assessment

The Heart

Section A *(6 marks)*

Label the parts of the heart

1. ______________________________

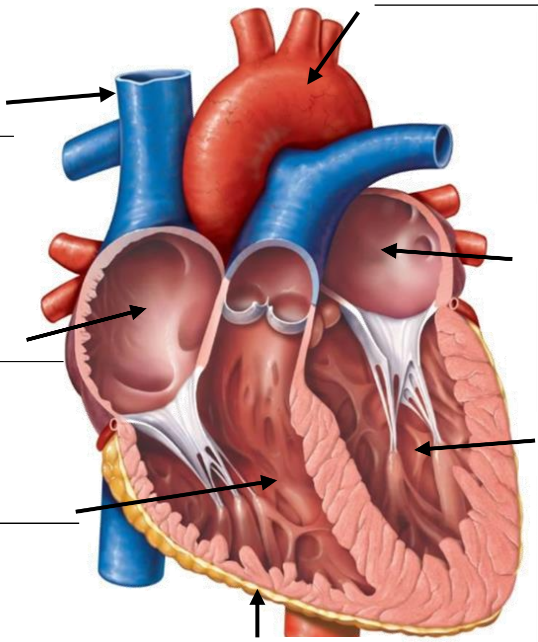
7. ____________________

2. ____________________

6. ________________________

3. ____________________

5. ________________________



4. ______________________________

Write as much as you know about how the heart works: *(3 marks)*

Section B *(5 marks)*

1. Thomas and Emma are good friends. Whose heart is larger? *(1 mark)*________
2. Aseel buys a 500g tub of butter. How much more is this than an adult heart? *(1 mark)*

   ____
3. Ryan was born with a damaged heart; one chamber does not work properly. How many chambers do work? *(1 mark)* _________________
4. Sarah's mum is sitting on a bench reading a book. Her heart rate is how many beats per minute? *(2 marks)* ______________________

Appendix F – Map Text-Only Condition

Map–Text (Paper, M-technology)

Vista Town

Vista is a friendly town that has a park in the centre where children go to play. The road that runs around most of the park is called **Vista Ave West** and **Vista Ave East**. The road that runs north of the park is called **Haight Street**. In the south of the park there is a **Zoo**. To the west of the park there is a **greenhouse** and just south of that, a **church**. To the south of the park there is a **castle**. East of the castle and south of the park is the **school.** There is a dead tree to the left of the school. There is a random **dinosaur** roaming the village, currently to the south east of the town. There is a children's **hospital** to the east of the town, and a **supermarket** on **Haight Street** to the north east, just after the park.

1. Sarah and John live on the corner of **Haight Street** and **Poppy Ave** which is in the north west of Vista Town. They have a puppy called Duke who ran away to find them at school. They need to go and fetch him, a 10 minute walk.

   From the corner of their house they turn right into **Poppy Ave** and walk south. One block down, they pass a **greenhouse** on their right. They continue along Poppy Ave. The next block down, they pass a **church** on their left. They continue along Poppy Ave. They walk another 2 blocks and the road turns in a south westerly direction, there is a **castle** on their left. They cross over a road (**Upper Lane**) and Poppy Ave changes to **Roose Way** going east. They walk one block, passing a **dead tree** on their left, and come to the school which is on their right. They find their puppy.

2. Sam has a guitar lesson in **Fred Street**, west of the park, and needs to go and meet his friends at the zoo afterwards. He comes out of the building and turns left. He walks along Fred Street for three blocks, passing the **church** on his left at the third block. He turns right into **Vista Ave West** and walks in a southerly direction for two blocks. The **castle** is on his right and the **Zoo entrance** on his left. He meets his friends.

Appendix G – Map Text, Graphics & Audio Condition

Map–Text, Graphics and Audio (Paper, M-technology)

Vista Town

Vista is a friendly town that has a park in the centre where children go to play. The roads around the park are called **Vista Ave West** and **Vista Ave East.** The road that runs north of the park is called **Haight Street**. In the south of the park there is a **Zoo.** To the west of the park there is a **greenhouse** and just south of that, a **church.**

To the south of the park there is a **castle.** East of the castle and south of the park is the **school.** There is a dead tree to the left of the school.

There is a random **dinosaur** roaming the village, currently to the south east of the town. There is a children's **hospital** to the east of the town, and a **supermarket** on **Haight Street** to the north, just after the park.

1. Sarah and John live on the corner of **Haight Street** and **Poppy Ave** which is in the north west of Vista Town. They have a puppy called Duke who ran away to find them at school. They need to go and fetch him, a 10 minute walk.

   From the corner of their house they turn right into **Poppy Ave** and walk south. One block down, they pass a **greenhouse** on their right. They continue along Poppy Ave. The next block down, they pass a **church** on their left. They walk another 2 blocks and the road turns in a south easterly direction, there is a **castle** on their left. They cross over a road to **Roose Way** going east. They walk one block, passing a **dead tree** on their left, and come to the school which is on their right. They find their puppy.

2. Sam has a guitar lesson in **Fred Street**, west of the park, and needs to go and meet his friends at the zoo afterwards. He comes out of the building and walks towards the park. He walks along Fred Street for three blocks, passing the **church** on his left at the third block. He turns right into **Vista Ave West** and walks in a southerly direction for two blocks. The **castle** is on his right and the **Zoo entrance** on his left. He meets his friends.

Appendix H – Map Assessment

Map Assessment

Quiz

Name: **Tutor Group**:

Section A *(12 marks)*

1 What is the name of the town? *(1 mark)*

2 There is a random animal walking around–what animal? *(1 mark)*

3 At the corner of which two streets do Sarah and John live? *(2 marks)*

4 Describe how Sarah and John walked to school to fetch their puppy. *(3 marks)*

5 What was Sam doing before meeting his friends at the zoo? *(1 mark)*

6 Describe how Sam would get to the zoo. *(2 marks)*

7 Write down anything else you can remember. *(2 marks)*

Section B *(2 marks)*

1 If Sam followed the road around the east of the park to the north, what would he see on his right? *(Question in the immediate recall testing time, 1 mark)*

2 How would Sarah and John walk to the supermarket from their house? *(Question in the immediate recall testing time, 1 mark)*

3 If Sarah & John walked north of the park, what would they see ahead? *(Question in the delayed recall testing time, 1 mark)*

4 If Sam followed that road around the east of the park to the north, what would he see on his right? *(Question in the delayed recall testing time, 1 mark)*

Appendix I – Poem Text-Only Condition

The Fox and Crow–Text (Paper, M-technology)

The Fox and the Crow

There once was a **crow** whose voice was loud and strong,
In his tree he would sit and talk, all day long.
One day the crow was hungry, his stomach growled so,
When suddenly he saw some **cheese** down below!

It was a yummy piece, so he flew down to the ground,
And snatched it before anyone else could come round.
He then quickly returned to his **branch** way up high,
To enjoy his little snack from his place near the sky.

A **fox** came round the corner, also in a hungry mood,
In fact, the fox had spent all morning looking for some food.
When he saw the cheese in the crow's mouth, he thought: "it's time to dine,
If the crow opens his mouth, the cheese will fall and then be mine."

"Hello there crow," the fox said, "how are you? Are you all right?"
But the crow did not respond – he held his beak shut very tight.
He thought: "poor fox! He's hungry and he's got nothing to eat,
And I have found this lovely cheese, my special yellow treat!"

"Crow, your wings are big and strong," the fox said with a grin,
"I bet you could fly anywhere, the shape your wings are in!
Not only do you have big wings, you've got that tasty cheese,
How did you get so lucky crow? Won't you tell me, please?"

The crow felt good to hear the fox, he flapped his wings with **pride**,
But still he didn't say a thing, he kept his words inside.
Although **he loved to talk**, this time he didn't speak,
He didn't want his yummy cheese to fall out of his beak.

"Hey crow," the fox said, he just wouldn't give in,
"That is one giant tree you're living in.
Not only are you smarter than all other crows,
You live in the highest tree anybody knows."

The crow felt really pleased, he felt very **proud**
But still he didn't say a word out loud.
He thought "**the fox must really like me**, after all I've heard,"
But he held onto his cheese, and didn't say a word.

"Just one last thing, dear crow, and then we will part,
I just wanted to tell you, from the bottom of my heart,
That **your voice is so lovely**, and more than anything,
It pleases me so to hear you sing."

The crow heard this praise, and beamed with pride,
He could no longer stay quiet – as hard as he tried.
He finally opened his mouth to sing a song,
Just like the fox had planned all along.
The cheese tumbled out from the mouth of the crow,
And fell straight to the fox that waited below.

The fox chewed it quickly, enjoying every bite,
"Thank you crow, my dear friend, it was a delight.
In return, I'll give you a word of advice,
When someone praises you so, you'd better think twice.
Praise is not something that's easy to earn,
And if someone gives it freely, he must want something in return."

The End.

Appendix J – Poem Text, Graphics & Audio Condition

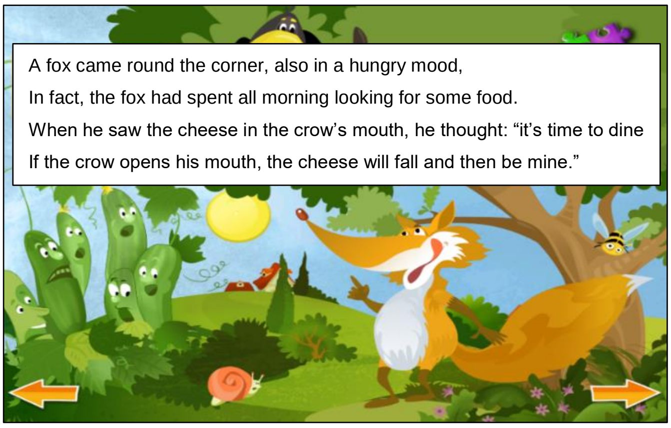
The Fox and the Crow–Text and Graphics (and Audio) M-technology

A sample of the layout for the M-technology text and graphics condition.

Appendix K – Poem Text & Graphics Paper

The Fox and the Crow, Text and Graphics (Paper)

A sample of the layout for the paper text and graphics condition.



Appendix L – Poem Assessment

The Fox and the Crow Assessment

Quiz

Section A *(7 marks)*

1 What food was lying on the ground? *(1 mark)*

2 What colour was the food? *(1 mark)*

3 Who got the food first? *(1 mark)*

4 Where did the animal take the food? *(2 marks)*

5 What is the moral (lesson) of the story–what can we learn from it? *(2 marks)*

Section B *(6 marks)*

1 What was the crow talented in (what was he good at according to himself)? *(1 mark)*

2 How did the other animal manage to get the food? *(2 marks)*

2 Why did the crow want to eat the food in the tree? *(1 mark)*

3 Who was more clever–the fox or the crow? Why? *(2 marks)*

Appendix M – Motivation Questionnaire

Motivated Strategies for Learning Questionnaire

Please rate the following statements based on **all** your subjects overall.

1 = not at all true of me to 7 = very true of me

Shade/underline the circle that is true for you

(If you make a mistake, put a line through it and shade/underline a new circle)

1 I like class work that is challenging so I can learn new things.

Not True True Very True
1 2 3 4 5 6 7

2 Compared with other students in my classes I expect to do well.

Not True True Very True
1 2 3 4 5 6 7

3 I have an uneasy, upset feeling when I take a test.

Not True True Very True
1 2 3 4 5 6 7

4 It is important for me to learn what is being taught in class.

Not True True Very True
1 2 3 4 5 6 7

5 I like what I learn in most subjects.

Not True True Very True
1 2 3 4 5 6 7

6 I can understand the ideas taught in most lessons.

Not True True Very True
1 2 3 4 5 6 7

7 I often choose topics I will learn something from even if they require more work.

Not True True Very True
1 2 3 4 5 6 7

8 I think that what I am learning in lessons is useful for me to know.

Not True True Very True
1 2 3 4 5 6 7

9 Compared with other students in this class I think I know a great deal about different subjects.

Not True True Very True
1 2 3 4 5 6 7

10 I worry a great deal about tests.

Not True True Very True
1 2 3 4 5 6 7

11 Understanding my subjects is important to me.

Not True True Very True
1 2 3 4 5 6 7

12 I ask myself questions to make sure I know the material I have been studying.

Not True True Very True
1 2 3 4 5 6 7

13 It is hard for me to decide what the main ideas are in what I read.

Not True True Very True
1 2 3 4 5 6 7

14 When work is hard I either give up or study only the easy parts.

Not True True Very True
1 2 3 4 5 6 7

15 I work on practice exercises and answer end of chapter questions even when I don't have to.

Not True True Very True
1 2 3 4 5 6 7

16 Even when study materials are dull and uninteresting, I keep working until I finish.

Not True True Very True
1 2 3 4 5 6 7

17 I work hard to get a good grade even when I don't like a class.

Not True True Very True
1 2 3 4 5 6 7

18 I find that when the teacher is talking I think of other things and don't really listen to what is being said.

Not True True Very True
1 2 3 4 5 6 7

Appendix N – Student Informed Consent

Initial Research Questionnaire

Initial Research Questionnaire

Name: **Tutor Group:**

Language you speak most at home: English

Please answer the questions honestly.

1 **What do you think about being asked to take part in a research project?** (Excited, worried/not worried, nervous, it doesn't affect you at all?)

I don't feel worried at all as it might be an exciting project

2 **Thinking about different ways of learning, what is the best way you learn?**

In my opinion, the best way I learn is by using interactive work.

3 **Do you agree to your information being used? You will be assigned a number so no one will be able to identify you – it will not change your grades.**

Underline the answer: <u>YES</u> NO

If you do not agree, you will still need to take part in the activity, but your information will not be used.

Thank You

☺

Appendix O – Students General Information

Post Research Questionnaire

Thank you very much for taking part in this study ☺

Please answer the questions honestly.

		The Heart	Vista Town Map	The Fox and The Crow
1	**Number the topics in order of preference 1, 2 and 3.** 1 being the best and 3 being the least liked *If you missed a session, just leave out that one.*	2	3	1
2	**Why did you enjoy number 1 the most?** I enjoyed the fox and the crow because it was an interesting story and easy to remember.			
3	**Why did you enjoy number 3 the least?** It was very hard and complicated and not easy to remember.			

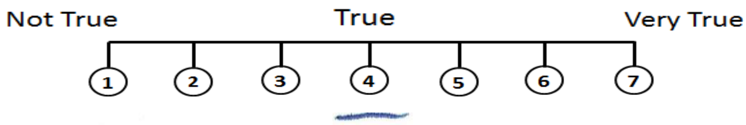
		The Heart	Vista Town Map	The Fox and The Crow
4	**How many times, approximately, did you go through the material for each topic?** Write how many times you think you read through or listened to the material.	3	10	2
5	**Was 5 minutes enough – or did you want more or less time?** I think it would be better if we had different amounts of time for different topics.			
6	**Were there any distractions going on around you that stopped you from learning the material? If so, what were they?** There weren't any distractions for me.			
7	**Did you spend all 5 minutes going through the material, or did your mind wander? Why?** I spent around 5 minutes but my mind was wandering off for a minute.			
8	**Was there enough time at the end to answer the questions properly? Would you have liked more/less time?** There was enough time to answer the questions because there weren't many.			

9	**Did you like using the iPad to study/learn from? Why/why not?** I like to study from iPads because it seems more interesting.			
10	**Answer A or B**			
	A If you had different topics using paper and iPad, which did you prefer – paper or iPad? Why? iPad because it is less boring.			
	B If you just used the iPad, which topic did you prefer? Why?			
11	**How do you think you learn and remember things best?** iPad because you will be able to picture and remember it in your brain.			
	How well did you understand the language?	**The Heart**	**Vista Town Map**	**The Fox and The Crow**
12	**Were there any topics where you did not understand some of the words?** Write **yes** under the topic heading if you did **not** understand some words.	No	Yes	Yes

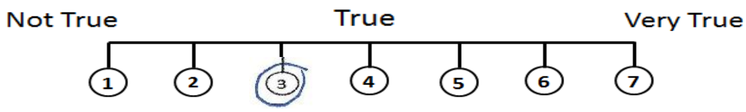
Once again, thank you very much ☺

Underline the correct answer.

1 I tried to complete this quiz to the best of my ability.



2 I am interested in the results of the research



Appendix P – Instruction to Assisting Teachers

Instructions to teachers

1 Ensure desks are spread out so students cannot see each other's' work.

2 Ask students to take out something to write with (pen/pencil – sharp) and their iPads (covers to remain **closed**).

3 Take the register – then SILENCE from now on.

4 Do you remember the research you agreed to take part in for Ms Thompson? We are going to do some today – the Maths department is helping with this.

5 You are going to be given a general quiz to see how much you know about a topic. This will take up to 5 minutes and must be completed in absolute **silence**. You may not ask questions during this time – I can't answer any of your questions. Please do **not** worry if you do not know anything – it is not a problem at all. Ms Thompson just wants to see if you know anything about this topic.

6 **Hand out the initial Quiz (1) sheets <u>now</u>** – face down. Ask students to turn over and write their name and tutor group - start timing. **Please be vigilant the whole time and check that students are not looking at other student's work**. **Collect papers – still NO TALKING**.

7 **Tell students** – I am now going to give you some information I would like you to study for 5 minutes – study it as you would if you were studying for a text – go over and over it. Afterwards we will carry on with our lesson as normal. Thank you. Still **no talking**.

8 **Ask students to open WebDavNav – go to their <u>tutor group folder – Quiz 1</u> and open the document in the Quiz 1 folder (press down on the document and 'open in … Pages'. If they have problems, they should raise their hands – they cannot help another student.**

9 Once all the students have opened it, start the 5 minute timing. At the end of 5 minutes, ask them to close the document and **DELETE** it **– this is important** .

10 **Resume your normal lesson – do not say anything more about the research, say you are not allowed to answer any questions -** Ms Thompson will be coming around, you can ask her any questions.

11 15 minutes before the end of the lesson students should be back in same position/location as at the start of the lesson.

12 **SILENCE** now. Reassure students – Ms Thompson just wants to see how much you remember naturally. This will not affect your grades and your parents will not be told how you do individually.

13 **Ask students to open WebDavNav – go to their <u>tutor group folder – Quiz 2</u> and open the document in the Quiz 2 folder (press down on the document and 'open in … Pages'. Complete the quiz, rename the document as their own name and upload into the Quiz 2 folder (you can track on your iPad) Please be vigilant the whole time and check that students are not looking at other student's work**. (instructions for students should be at the bottom of the Quiz 2 document). They will need as much time as possible.

14 Tell students: Thank you for taking part. Please do not worry if you think you have not done very well – this is not a problem, it is to help **us** learn how to teach better.

Appendix Q – Mode of Presentation Language Analysis

Analyses of variance determined one main effect of language in the heart topic, and no interactions in any of the three curriculum topics.

ANCOVA Main Effects and Interactions for Mode of Presentation (Paper, M-technology) for Each Curriculum Topic (Heart, Map, Poem)

Topic	Source	*Mean Square*	*df*	*MSerror*	*F*
Heart	Language	58.14	1, 206	7.91	7.35 **
	Mode * Language	20.22	1, 206	7.91	2.56
	Time * Language	19.37	1, 206	4.12	4.70
	Time * Mode * Language	11.00	1, 206	4.12	2.67
Map	Language	3.56	1, 190	6.97	.51
	Mode * Language	0.04	1, 190	6.97	.01
	Time * Language	0.32	1, 190	2.45	.13
	Time * Mode * Language	0.01	1, 190	2.45	.00
Poem	Language	4.19	1, 207	5.24	.80
	Mode * Language	1.16	1, 207	5.24	.22
	Time * Language	0.10	1, 207	2.47	.04
	Time * Mode * Language	0.33	1, 207	2.47	.13

Note. *$p < .05$. **$p < .01$. ***$p < .001$.

Heart – language (English –M = 8.17, SE = 0.17; Other M = 7.38, SE = 0.24)

Appendix R – Cognitive Load Language Analysis

Analyses of variance determined no significant main effects or interactions in any of the three curriculum topics including language.

ANCOVA Main Effects and Interactions for Cognitive Load (Text-Only Text & Graphics; Graphics & Audio; Text, Graphics & Audio) for Each Curriculum Topic (Heart, Map, Poem)

Topic	Source	*Mean Square*	*df*	*MSerror*	*F*
Heart	Language	8.43	1, 194	7.88	1.07
	Load * Language	0.69	3, 194	7.88	.08
	Time * Language	9.80	1, 194	4.40	2.23
	Time * Load * Language	7.99	3, 194	4.40	1.81
Map	Language	.05	1, 190	8.29	.01
	Load * Language	1.54	3, 190	8.29	.19
	Time * Language	3.19	1, 190	2.96	1.08
	Time * Load * Language	0.93	3, 190	2.96	.31
Poem	Language	.33	1, 199	4.83	.07
	Load * Language	6.53	3, 199	4.83	1.35
	Time * Language	.01	1, 199	1.90	1.81
	Time * Load * Language	3.43	3, 199	1.90	1.81

Note. *p < .05. **p < .01. ***p < .001.