

LEVEL 7 ASSESSMENT SPECIFICATION

Student name:	Katherine Cicharski-Harwood	Student ID number:	1800488
Programme:	Secondary PGCE		
Module:	Curriculum Development and Enhancement		
Module code:	ECSP7002	Contribution to Overall Module Assessment (%):	100%
Lecturer:	Helen Denny	Internal Verifier:	
Assignment Title:	Research Study including Ethics Form	Word count (or equivalent):	10.000 words in total including an Ethics form (to be given in on January 11 th 2019)
Submission deadline:	May 13 th 2019	Return date of provisional marks & written feedback:	June 10 th 2019
Submission method:	All written assessments, where practical and possible, must be submitted via Turnitin unless otherwise instructed by the Lecturer. (Please DO NOT put this assessment specification into Turnitin or it will match many similarities with other students' submissions.) Alternative submission method (if applicable): Late submission of the assessment will result in a late penalty mark. Penalties for late submission: Up to one week late, maximum mark of 50%. Over one week late, 0%. Only the Extenuating Circumstances Panel may approve a change to submission dates.		
Academic honesty / referencing:	Academic honesty is required. In the main body of your submission you must give credit to authors on whose research and ideas your work is based. Append to your submission a reference list that indicates the books, articles, etc. that you have used, cited or quoted in order to complete this assessment.		

How effective are problem solving, ‘desirable difficulties,’ ‘cognitive conflict’ and social interaction in committing new vocabulary learning to long term memory?

Introduction

In accordance with the British Educational Research Association’s (BERA) Ethical Guidelines for Educational Research, the school in which I carried out this research during my first professional teaching experience (PTE1) will remain anonymous and will be referred to as School X (British Educational Research Association, 2011). The school, with, at present, 732 pupils on roll, is located in a sector that the 2015 Estyn report characterized as ‘deprived’ with around 60% of pupils living in the 20% most disadvantaged sectors in Wales (Estyn report, 2015). Moreover the percentage of pupils with special educational needs is around 37%, much higher than the national average of 25%. Despite these disadvantages, Estyn rated the school’s current performance as ‘excellent.’ For my research, I worked with a top band class from Year 7, composed of 32 pupils, of which 15 were considered MAT (more able and talented) pupils.

My research activity is based on vocabulary acquisition, focusing on how to commit new learning to long-term memory, adopting a cognitive science approach. This approach, summarized in *Deans for Impact*, suggests that thinking about the meaning, and practising new information, helps it to be stored and later retrieved (Deans for Impact, 2015). Moreover, as not all practice is equal in value (Ericsson *et al.*, 1993), practising, enhanced by ‘desirable difficulties,’ such as ‘inter-leaving’ and testing (Bjork, 1994), introducing ‘cognitive conflict’ (Piaget 1975), creates difficulties which require more effort on the part of the learners to overcome. The effort of trying to remember creates more durable and recallable memories (Agarwal *et al.*, 2012). Furthermore, if such difficulties are surmounted when working in social interaction with their peers, the information is then internalised into their long-term memory (Vygotsky, 1987).

Literature Review

Over the past century, various theories have been elaborated to explain the learning process and how best to achieve long-term memorisation and learning. Skinner’s Behaviourism in the 1930s conditioned learners by rewarding or ‘punishing’ them for their learning outcomes. Piaget (1896-1980) asserted that learning occurs in stages in a child’s development as a result of internal processes which the child constructs to make sense of the world. Vygotsky (1896-1934) added to this constructivist theory by arguing that the learner brings meaning to his/her world via social interaction: this meaning is then internalised on an individual level by the use of language. Vygotsky also argued

that there existed a zone of proximal development, a gap between what the learner can do by him/herself and what he/she could do with help, or scaffolding.

However, it is only since the turn of the 20th/21st century that brain research and neuroscience have developed sufficiently to confirm or refute these theories. In fact it has done neither, rather it has added a new dimension to our knowledge. Cognitive science has, above all, revealed the crucial importance of memory in the learning process. 'Memory' has been found to be a process itself, composed of three parts: a sensory input buffer (sight, hearing, touch, taste, smell) which is very transitory, lasting from 0.25-2 seconds. This sensory input is then converted to be used in the working memory and long-term memory. The working memory uses information from both the sensory buffer and the long-term memory: it is where we 'think'. Information can only be stored there for around twenty to thirty seconds and its capacity is limited. Long-term memory is made up of a neural network with nodes representing chunks, or concepts, connected by links. The nodes and links form schemata, clusters of information. When one part is activated, the whole cluster becomes active (Best, 2000). Thus, according to the cognitive information processing theory, our purpose of recording new vocabulary into long-term memory and then being able to make connections to retrieve it, requires the use of two discrete and 'crucial components' of learning (Muijs, 2010).

Having discussed the manner in which these two functions (memorisation into long-term memory and creating connections to retrieve it) occur in the learner's brain, we realise that a teacher's aim is to teach in the most effective way to help these functions happen. Brain research reveals that the brain works best when meaningfully organising random information, such as during problem solving (Lackey, 1999), and also when the learning is connected to physical experience: we learn when we put the new knowledge into practice, 'we learn by doing' (Muijs, 2010). Brain research also informs us that students learn new ideas by reference to ideas that they already know (Bransford *et al.*, 2000), but warns against cognitive overload (Sweller, 1988) which can be reduced by the use of 'worked examples' (Paschler *et al.*, 2007) and by avoiding dividing the learner's attention (Chandler *et al.*, 1992).

The scope of this present research is therefore based on the cognitive principle that for information to enter the long-term memory it should not only be learnt in reference to ideas already known, but also, to be correctly encoded, it has to become relevant to the learner by his/her own effort, for instance by meaningfully organising random information in a practical, physical way. McDaniel *et al.* in their 1994 research '*What Makes Folk Tales Unique*' describe three such experiments. In the first, texts were presented in which the sentences were randomly ordered and the subjects had to reorder the sentences into a coherent text. In the second test the subjects were required to construct a meaningful outline for the text and in the third, the subjects had to fill in blanks in the text, as 18-30% of the letters had been deleted. A control group had simply read and re-read the text. In the subsequent recall tests it

was found that the information encoding task of unscrambling the randomly ordered sentences improved recollection significantly, in relation to the control group who had simply read the text. The task requiring subjects to construct a meaningful outline for the text also significantly enhanced recall. The other encoding task of filling in the blank letters also increased recall efficiency compared with reading the text, but less than the unscrambling and outline tasks. This suggests that filling in blank letters requires less deep semantic processing (McDaniel *et al.*, 1994). Another form of effort deployed by the learner in organising information and thus encoding it, is by requiring them to supply explanations concerning the information using ‘deep questions’ such as ‘why, why-not, how, what-if, how does X compare to Y, and what is the evidence for X?’ (Pashler *et al.*, 2007). This method correlates with asking the learners to construct a meaningful outline for a text in McDaniel *et al.*’s example above. Inventing stories and mnemonic imagery are also useful and effective in encoding information into long-term memory (Peters *et al.*, 1986). Levin, applied this strategy to the ‘keyword method’ of vocabulary learning (Atkinson, 1975). The ‘keyword’ is a familiar word that is similar to the unfamiliar, foreign word (Peters *et al.*, 1986). In the case described, the Spanish word *carta* (postal letter) could be recorded into the familiar English word *cart*. The keyword can then be related to the meaning via an image of a letter in a cart. Mnemonic images create a ‘systematic retrieval route’ (Peters *et al.*, 1986).

The learner’s efforts may also be channelled through various ‘desirable difficulties’ (Bjork & Bjork, 2011) in which conditions of learning are created that challenge the learner and slow down the rate of encoding. This creates optimal long-term retention and transfer (Bjork & Bjork, 2011). These ‘desirable difficulties’ are concentrated during the ‘practising’ of the new information. As we have seen above, learning is connected to physical experience: ‘we learn by doing’ (Muijs, 2010) and for information to be encoded and stored it has to be practised in many complex and interactive experiences (Muijs, 2010). In *The Role of Deliberate Practice in the Acquisition of Expert Performance*, Ericsson *et al.* state that the repercussion of practice on the optimisation of performance is greater than previously believed possible (Ericsson *et al.*, 1993) and *Deans for Impact* assert that practice is an essential part of learning new facts (Deans for Impact, 2015). However Bjork and Bjork argue that practice and training may improve performance without necessarily increasing learning: as stated in *Deans for Impact* ‘not all practice is equivalent’ (Deans for Impact, 2015). Practice involving cognitive effort is required for memory to be properly encoded. Indeed, Bjork and Bjork contend that learning may become contextualised if practice occurs under predictable conditions, and although information is initially easily retrieved, the recall performance of this information is poor at a later date. On the other hand modifying the context, or even the environment, enhances retrieval on a later test (Bjork & Bjork, 2011). They present the case of children throwing beanbags at a target. After a delay, those who practised throwing from a fixed distance did much worse than children who had been

practising from varying distances, although, in the final test, they all threw from the same (original) fixed distance. Variation during their learning was thus proved beneficial.

Bjork and Bjork also advocate the use of distributed practice. Mass practice such as ‘cramming’ presents good results for exams the following day, but not for long-term retention, whereas ‘spaced’ practice has been proved to bring results in the long term. Cepeda *et al.*, in their *Review of the Distributed Practice Effect*, differentiate the Interstudy Interval (ISI), the time separating two study episodes of the same material in the same study time (lesson), and the ‘lag effect,’ referring to the period between two learning episodes (lessons). However, Cepeda *et al.*, conclude that ‘spaced’ (vs. massed) learning consistently provides benefits, whatever the interval of retention, furthermore, learning benefits develop with larger time lags (Cepeda *et al.*, 2006). Bahrick *et al.* made a five-year research on the spacing effect on foreign language vocabulary retrieval. Their results demonstrated that spacing the learning intervals by two months, or possibly longer, reinforces long-term memory as they revealed extensive gains in vocabulary retention over five years (Bahrick *et al.*, 1993). Bjork and Bjork affirm that spacing has demonstrated some of the most common and powerful results in the experience of experimental research on learning and memory. The IES Practice Guide recommends teachers to expose pupils to key concepts on at least two occasions, distanced by a week to several months. Homework, quizzes and exams should also be delayed from the first learning session. (Pashler *et al.*, 2007). Spacing practice has the added advantage of providing a foundation for new learning which can be built onto this practice of past (prior) learning in the knowledge that students learn new ideas ‘by reference to ideas they already know’ (Deans for Impact, 2015).

A similar ‘desirable difficulty’ advocated by Bjork and Bjork is ‘interleaving’ (alternating practice) by alternating topics. Switching between topics can also be a way of introducing spacing. In an experiment described by Bjork and Bjork, participants were given eighteen trials on a pinball-like apparatus using three different movement patterns. In the interleaved group, practice was arbitrary whereas in the ‘blocked’ group, each movement pattern was practised separately. Participants in the ‘blocked’ group improved quicker than the interleaved group. However, in random testing ten days later, the interleaved participants far out-performed the ‘blocked’ group and even in blocked testing, interleaved participants were still better performing. Blocked practice may thus appear to be the most advantageous for initial learning but interleaved practice develops long-term memory, transfer of skills and better future retrieval. Bjork and Bjork suggest that when learners have to manipulate the interference between the differing subjects under study, they are compelled to remark similarities and disparities between them. This results in the encoding of higher-order models which, in turn, promotes both retention and recall. A second theory contends that learners are obliged to reload and thus recode memories when interleaving. Repeated reloadings would promote learning and memory retrieval at a later date (Bjork & Bjork, 2011). Another study, carried out by Rohrer *et al.* in 2015, compared interleaved mathematics practice with the usual blocked approach. In this study, portions of previous

courses were practised in an interleaved order, obliging students to choose a strategy based on the problem itself. Tests given, one or thirty days later, produced considerably higher scores for interleaved practice. Rohrer's study was also the first to show that the increase in positive results, permitted by interleaving, does not decline with time, rather, it grows larger, providing a quasi-invulnerability against forgetting (Rohrer *et al.*, 2015).

The effort of retrieval is perhaps the most potent learning effort or 'desirable difficulty' in obtaining long-lasting memorisation. *Deans for Impact* asserts that trying to remember something makes memory stronger and more durable than other forms of studying (Deans for Impact, 2015). As long as 2300 years ago, Aristotle wrote 'Exercise in repeatedly recalling a thing, strengthens the memory' (Aristotle, circa 350 BC). Empirical research has since demonstrated that repeated retrieval enhances long-term retention more than studying (Karpicke *et al.*, 2008 / Carpenter *et al.*, 2008 / Kuo *et al.*, 1996 / Roediger *et al.*, 2006 / Toppino *et al.*, 2009 and many others). Indeed, Roediger and Butler's study on *The critical role of retrieval practice in long-term retention* (2011), draws several interesting conclusions: as well as finding that retrieval practice regularly produces greater long-term retention than studying, they assert that repeating the tests obtains better results than single testing. Feedback after testing further improves the results, with delayed feedback (given after the test) being even more effective (perhaps due to the spacing effect). Roediger's research also shows that a lag of time between study and test is necessary for the retrieval practice to be effective (wide spacing of retrieval attempts generate the best results). Furthermore, the 'mnemonic benefits' of testing generate transferable and flexible knowledge that can be used in different contexts, such as applying the knowledge to different situations. In all of these retrieval practices, 'easy conditions' of retrieval (such as testing immediately after study), produce little benefit for long-term retention, whereas 'effortful processing' in more 'difficult' conditions, including spacing, interleaving and repeating the tests, present real benefits which remain over a long period of time (several years), (Roediger *et al.*, 2011).

Tests, for many years considered as neutral assessment tools, should thus be integrated into educational practice to boost performance. They can be introduced in various ways such as self-testing with flash-cards, end-of-chapter questions, or quizzes/tests (Roediger *et al.*, 2011). Furthermore 64% of students tested with quizzes reported that they felt that their learning had increased and that their test anxiety before end-of-term exams had decreased (Agarwal *et al.*, 2012). Karpicke *et al.* affirm that pre-tests, however, such as quizzes carried out before a lesson, do not enhance results at end-of-term exams compared with no-quizzing (Karpicke *et al.*, 2012), yet Pashler *et al.* recommend using pre-tests to introduce new topics so that learners can identify what points they do not yet know (Pashler *et al.*, 2007).

Theories, to explain how the retrieval practice effects work, include the notion that retrieval of information produces amplification of the memory trace, making it easier to be retrieved again in the

future. Another theory centres on the effort involved in retrieval: more effort would induce more extensive reprocessing. Bjork and Bjork distinguish between storage strength (permanence) and retrieval strength (accessibility). Their theory proposes that once an information with weak storage strength is retrieved, its retrieval strength will be reinforced for some time afterward (Roediger *et al.*, 2011). Their basic message is that we should spend more time testing ourselves and less time ‘re-studying.’ We should not forget that retrieval/testing modifies our memory by transforming the information recalled to become, not only more likely to be retrievable in the future, but also retrievable in varying situations (Bjork & Bjork, 2011). As Bjork and Bjork argue, ‘generating’ a solution, rather than looking up the answer or being told the correct response, creates a long-term benefit in learning. When we make use of past knowledge and existing clues to generate our own answer to a problem, we develop our long-term memory of the information: retrieval is, itself, an impressive ‘memory modifier’ (Bjork, 1975).

Cognitive conflict is also a form of cognitive difficulty that changes the concepts of learners when new information does not fit in with prior knowledge. For Piaget, this conflict is overcome by individual activity whereas for Vygotsky, such difficulties are resolved in interaction with others, through collaboration in pair or groupwork. For Vygotsky, social interaction, such as in problem solving in pair or groupwork, is essential for information to be committed to long-term memory, cognitive functions having their origin in social interaction that is subsequently internalised (Vygotsky, 1987). Socio-cognitive conflict has been used in activities with learners to change their concepts on various scientific concepts (Skoumios, 2009) or to enhance rational teaching (Sacco *et al.*, 2008). Mogonea *et al.*’s research, in a study involving one hundred and twenty-three student teachers, concludes that the process of solving cognitive conflicts into meaningful concepts, through interaction with others, has major benefits on the development of meta-cognitive capabilities (Mogonea *et al.*, 2014).

My research activity, based on how to commit new learning (vocabulary acquisition) to long-term memory, thus adopted ‘desirable difficulties’ and problem-solving to be resolved in social interaction in pair and groupwork, to foster long-term memorisation.

School Policy and Practice

As suggested in the introduction above, School X’s whole ethos is based on giving each pupil a top standard education, by differentiating its teaching and challenging each learner to achieve his/her maximum capabilities, the school’s motto being “If you believe it, you can achieve it!” Within a ‘deprived,’ low ambition, and inevitably low achievement environment, the school is fighting for each pupil to believe in him/herself, stretch upwards, aspire, and take confident pride in their achievements.

However, Modern Foreign Languages, as a subject, is in the invidious position of neglect throughout the United Kingdom with a 29% fall in GCSE language entries over the past five years in Wales (and an up to 50% fall in England). Schools throughout the country indicate that the reason for this drop is the perception of languages as a difficult subject.¹ In these problematic circumstances, making the learning of foreign languages more of an ‘effort’ and creating ‘desirable difficulties’ with the intention of assisting and enhancing second language memory storage and facilitating language retrieval might add to the presumed ‘difficulty’ and undesirability of the subject, particularly as ‘Modern Foreign Languages’ is an optional subject after KS3 so very easily dropped.

In the Modern Foreign Language department of School X, there is no active policy of deliberately creating supplementary and arbitrary challenges in learning with the avowed intention of optimising long-term retention and memory transfer. The scheme of work and department resources provide wide variation in the learning tasks and puzzles such as mixed up words, missing letters or missing spaces; and very occasional tests are scheduled, mainly for the end of term. Each School Year has its MFL (French or Spanish) Booklet, a copy of which is given to each language student. It comprises glossaries of all the vocabulary used, and lists of all the structures studied, including verbs, grammar, key phrases etc. There are also short comprehension texts and exercises such as sentences with missing words to fill in. The teachers have a ‘Teacher Handbook’ uniting all the Schemes of Work for each Pupil Year and the school Assessment Policy. Within the MFL teaching department there is free access to PowerPoint lessons which can be used if a teacher does not make his/her own resources for a lesson. Assessment criteria for homework, end-of-term assessments and GCSE requirements are also provided.

In the department, emphasis is placed on creating excellent differentiation to challenge each pupil to achieve his/her maximum possibility. Lessons are to be clear and varied with plenty of pairwork and motivating tasks and games. Teachers enthusiastically assist, compensate difficulties and motivate the learners. However systematic testing, spacing, interleaving, deep questions and cognitive conflict do not consciously form an integral part of this programme.

As my mini-research is thus distinctively different from the way ‘normal’ lessons are taught in School X, an analysis between my research lessons and results with ‘normal’ outcomes may be considered interpretable, despite the fact that there will be no objective quantitative comparison as there is no control group.

¹ ‘*Language Trends 2018. Language teaching in primary and secondary schools in England*’ Survey Report by Teresa Tinsley and Neela Dolezal. www.britishcouncil.org

Methodology & Ethics

The design frame of this study is an action research, that is, a research initiated by practitioners such as teachers, health or social workers, with the intention of developing their practice or dealing with a problem (Thomas, 2013). An action research is a problem-solving enquiry ‘conducted by and for those taking the action’ in order to help improve and perfect their own activity (Impact Wales, 2018). This principle of action research was initiated by the social psychologist, Kurt Lewin in his 1946 paper ‘*Action research and minority problems*’ in which he argued that it is not sufficient for research to simply produce books, it has to lead to action. For Lewin, action research should bring about *action*, based on events composing a ‘circle’ of planning and action, followed by fact-finding about the results of this action (Lewin, 1946). The ‘circle’ should continuously spiral upwards: reflection on the results of action, and changes made, should stimulate further questions and ideas, which again can be put into action, and the results would then be observed and reflected upon. It thus builds upon the practitioner’s discoveries in a process of ‘planning, reflection and re-planning’ (Thomas, 2013).

Researchers exploit a collection of assumptions to attempt to understand and interpret research material and transfer it into functional data. These collections of assumptions or models/traditions are known as *paradigms* which have been described as an approach to research that contributes a structure consolidating perceptions of knowledge (Somekh *et al.*, 2011). The educational research paradigm adopted for my research is ‘interpretive,’ that is it attempts to interpret the world that it studies, in the knowledge that reality is a concept that we all comprehend in differing ways (Briggs *et al.*, 2012). Interpretive researchers perceive themselves as part of the research topic they study and recognise that the researchers are shaped by their participants in the same way as the researchers’ study affects the participants (Briggs *et al.*, 2012). As interpretive methodology does not consider there to exist an objective reality, research data is generally analysed qualitatively rather than quantitatively with a stress rather on words than numbers (Briggs *et al.*, 2012). However, Foskett *et al.*, advocates a combining of approaches following the needs of research, using a whole range of methodologies (Foskett *et al.*, 2005). A combined approach associating quantitative and qualitative procedures can balance the shortcomings of both.

The methodology of the present research is thus an interpretivist paradigm, aiming to observe and interpret the results of an analysis of the effect on memory, when teaching French vocabulary, of deliberately creating supplementary challenges ‘desirable difficulties’ such as problem-solving, puzzles, making sense of jumbled sentences and random information, spacing and interleaving; cognitive conflict; and problem solving, resolved by social interaction tasks. Due to the nature of the study results, obtained through testing, the research approach is mainly quantitative although qualitative elements will be used. My research thus focuses on long-term memorisation of a series of

French words, tested ten days after initial learning. After initial reflection on the preliminary results, it could be further developed, for example, by re-testing at spaced intervals over a much longer period, or with pupils of a different age and ability.

The subjects/participants of the present case study are a Year 7 class, learning vocabulary about hobbies. The case study will act as a template through which to examine the object of the research: the effectiveness of problem solving, ‘desirable difficulties,’ cognitive conflict and social interaction in committing new vocabulary to long-term memory. It is a single case study as it contains only one element, one class: my initial intention of comparing the study class with a control class had to be abandoned as my university, for ethical reasons, preferred not to make a comparison between class results. The study is bounded by time in a ‘snapshot’ manner (Thomas, 2013): a defined period of time will be studied, in this case two lessons, within an interval of ten days. Although this study is a single case, it is an example which can be extrapolated to all classes in which modern foreign language vocabulary is being taught. It is, in a certain manner, a ‘nested’ study acquiring its universality from the broad scope of its perspective (Thomas, 2013). My own role, in this interpretivist frame, is of a student-teacher researching the effectiveness of cognitive science theories on practical class-room teaching in a teaching environment that currently does not focus on these theories. My own positionality is of hopeful optimism, awaiting the results of this study to discern whether I should integrate these theories into my own practice.

The research tools used to examine the study question depend on the basis of the study itself, as with all the studies made on the topic of memory retention and retrieval, the preferred tool adopted is retrieval testing. The results of this retrieval testing can be compared between the two periods of testing (initial testing at the end of the first lesson, and ten days later at the beginning and end of the second lesson). The limitation of this tool and the methodology adopted is that there is no direct comparison of the quantitative results with a control group, a class of similar age and level, learning the vocabulary under School X’s habitual conditions.

Research in the classroom necessarily involves engagement with young and vulnerable people and so requires ethical scrutiny. Ethical considerations had already been announced and determined in the ethics form established in the initial stages of this research. However, when announcing the research to my research participant class, I also explained to them that they would all remain completely anonymous as would any material they gave in to me. Also any marking/marks arising from their work would in no way affect the school marking system or their grades. I thanked them in advance for their participation and wished them luck with all the testing. Globally the pupils felt motivated by the research and confident that it would not adversely affect them personally in any way.

Evaluation

Although, initially, I had planned to spread my study over a three lesson period, for practical, logistic reasons in the internal organisation of School X, this had to be revised down to two lesson periods. The two lessons were spaced at a ten-day interval and the participant class was a top band Year 7 class composed of thirty-two pupils with an equal number of girls and boys, three pupils with English as an Additional Language and three pupils receiving Free School Meals. The research had already been announced to the participant class and the students were aware of, and motivated to take part in, the study.

The spacing of ten days between the lessons corresponds to the ‘desirable difficulty’ of spacing, or distributing practice. As discussed above, Pashler *et al.*, in their *IES Practice Guide*, recommend space learning over time, from weeks to several months, with the resulting ‘delayed re-exposure’ leading to an increase in information remembered by the students (Pashler *et al.*, 2007). Moreover, the lessons of my research included testing which was therefore also spaced by this lag of ten days. Unfortunately, the time available for my research did not allow for a two-month interval between vocabulary retrieval, as recommended by Bahrick *et al.*, whose study on vocabulary retrieval revealed that leaving intervals of two months, or longer, between testing, demonstrated important benefits over the five-year study period (Bahrick *et al.*, 1993). However my ten-day spacing did allow for an interleaving effect between lessons, as during an intermediary lesson with this class, I taught a different topic, switching topics being one of Bjork & Bjork’s ‘desirable difficulties’ that compels learners to distinguish similarities and disparities between topics and reload and recode memories at each change of topic. The spacing also permitted the return to the initial topic of hobby vocabulary as prior knowledge, on which they could build further new ideas in reference to existing knowledge as affirmed in *Deans for Impact* (Deans for Impact, 2015).

The lessons were based on cognitive science methods, the ‘desirable difficulties’ advocated by Bjork and Bjork and the benefits of social interaction, working in pairs or groups (Vygotsky, 1987). In the first lesson, I introduced the vocabulary in the normal way, i.e., announcing aloud the word displayed on a slide with an image of its meaning, and drilling each word with the participants so that the pupils were used to the sound and pronunciation. After this traditional approach, I then gave the students a series of practical tasks, practice being ‘essential to learning new facts’ (Deans for Impact, 2015). These tasks all represented a form of difficulty or problem solving where considerable effort was required to make sense of, and organise, random information, given that brain research demonstrates that our brains are particularly good at organising chaotic data and then encoding the corrected messages effectively (Lackey, 1999). As stated earlier, these tasks were undertaken in pairs, so as to profit from their social interaction (Vygotsky, 1987). To avoid cognitive overload with these

problem solving tasks, worked examples were given on a systematic basis as recommended by Pashler (Pashler *et al.*, 2007).

The tasks included matching the French word with the image of its meaning, a form of retrieval practice of the new information which had just been announced, the interstudy interval (ISI) for this retrieval being around five minutes (Cepeda *et al.*, 2006). The participants then had to ask each other questions using prior knowledge from the previous lesson (thus focusing on a different topic, interleaving) along with the new vocabulary: in other words they had to make meaningful sense out of the new information, such as recommended by McDaniel *et al.* (McDaniel *et al.*, 1994). They then had to solve word puzzles such as word-searches for the new vocabulary, generating mental effort to recognise the new words. Subsequently they had to separate jumbled words to make coherent sentences and afterward translate the sentences, providing a meaningful signification and thus organising the information, again creating order out of chaotic information and encoding it (Lackey, 1999). Each re-utilisation of a new vocabulary word was spaced in interstudy intervals (ISI) (Cepeda *et al.*, 2006), of around five-ten minutes, and the words were reviewed each time in different ‘varied’ circumstances and new conditions, favourable to memory coding (Bjork & Bjork, 2011). All tasks were carried out as pairwork to promote social interaction. The first lesson ended with a test on all the new vocabulary, not essentially to assess the learning of the participants, but principally as an effort of retrieval, as discussed above, one of the most potent of the ‘desirable difficulties’ (Deans for Impact, 2015).

The second research lesson, after the ten day spacing, started with another test of the hobby vocabulary, both as an assessment to gauge the retrieval results following the first lesson and spacing, and as an effort of retrieval to re-code the information in the pupils’ memories. The lesson continued with interleaving to review learning undertaken in the intermediary lesson, again, at each switching of subject, obliging the students to reload and recode memories (Bjork & Bjork, 2011). It progressed on to further problem solving pairwork practice, including games of guessing the disappearing hobby, puzzles separating words to find letters hidden between them and which together form a further mystery word from the hobby vocabulary, and gap-filling of missing letters. These tasks were inspired by those set by McDaniel *et al.*, in their 1994 research. Further problem solving included arranging the words in alphabetical order and translation from English into French. The participants then undertook an oral conversation practice of the vocabulary with ‘we learn by doing’ (Muijs, 2010) in mind. Most of the tasks in the lesson were focused on the meaningful organisation of random information, creating order out of random data and encoding it (Lackey, 1999). All the tasks were undertaken in pairs or groups so as to benefit from social interaction. As in the first lesson, the use of hobby vocabulary was spaced in ISI (interstudy intervals) of around five to ten minutes in varied circumstances (Bjork & Bjork, 2011). The lesson ended with another test on the hobby vocabulary, again both as an assessment for my research results and as an effort of retrieval for the participants.

The lessons, tasks, social interaction and retrieval tests were carried out successfully and the retrieval results duly collected. An analysis of the three results of retrieval testing: testing at the end of the first lesson, testing ten days later at the beginning of the second lesson and testing at the end of the second lesson, will form the basis of an evaluation of the effects of problem solving and social interaction in hobby vocabulary acquisition, retention and retrieval.

However, the weakness of this methodology is evident: there is no direct comparison with a control class of similar aptitude taught using methods which do not include such testing and added ‘desirable difficulties.’ The results of my study, therefore, will not have an objective visibility as to the success of the research premise that problem solving, ‘desirable difficulties,’ cognitive conflict and social interaction are more efficient in committing new vocabulary learning to long term memory than traditional methods of teaching vocabulary. The comparison of testing results will therefore only demonstrate extremely good retrieval test results, middle-range results or mediocre ones. Any interpretation of such results risks becoming subjective and redundant as only a comparison with a control group or, possibly, retrieval tests taken with the participant group at a much later date, such as six months later, could afford conclusive, applicable quantitative data. Qualitative data, such as a survey or questionnaire addressed to the participant pupils could provide an idea of how the learners consider the strength of their learning but this would also be entirely subjective.

Given this structural weakness in the methodology, interpretative analysis must be undertaken with great caution and circumspection. An interpretation of the comparison between habitual results based on the usual teaching methods adopted by School X and the results of the mini-research must be implemented comparing very similar sets of vocabulary taught during a similar time period and to pupils of a similar age and level.

The results of the mini-research retrieval tests are, however, unambiguous. At the end of the first lesson, after the problem solving effected in pairs so as to profit from social interaction, the participants’ retrieval rate was 90% correct answers with a mean of twenty-two words correctly recalled out of twenty-four (see Figures 1 and 2), the median was twenty-two and a half words, the upper quartile being the full twenty-four words and the lower quartile, twenty words, with the lowest result being seventeen words correctly recalled (see Figure 3).

Figure 1

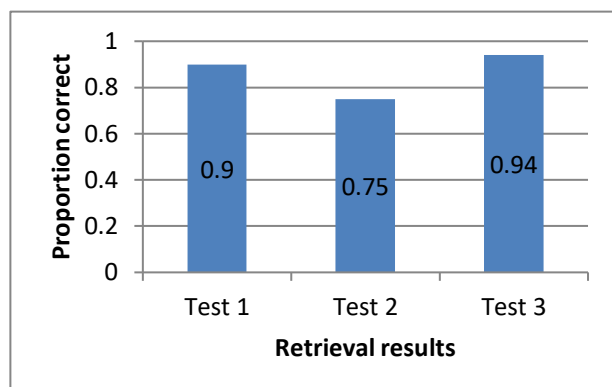


Figure 2

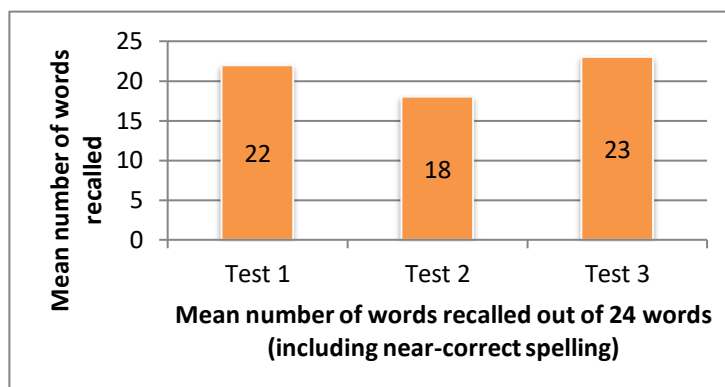
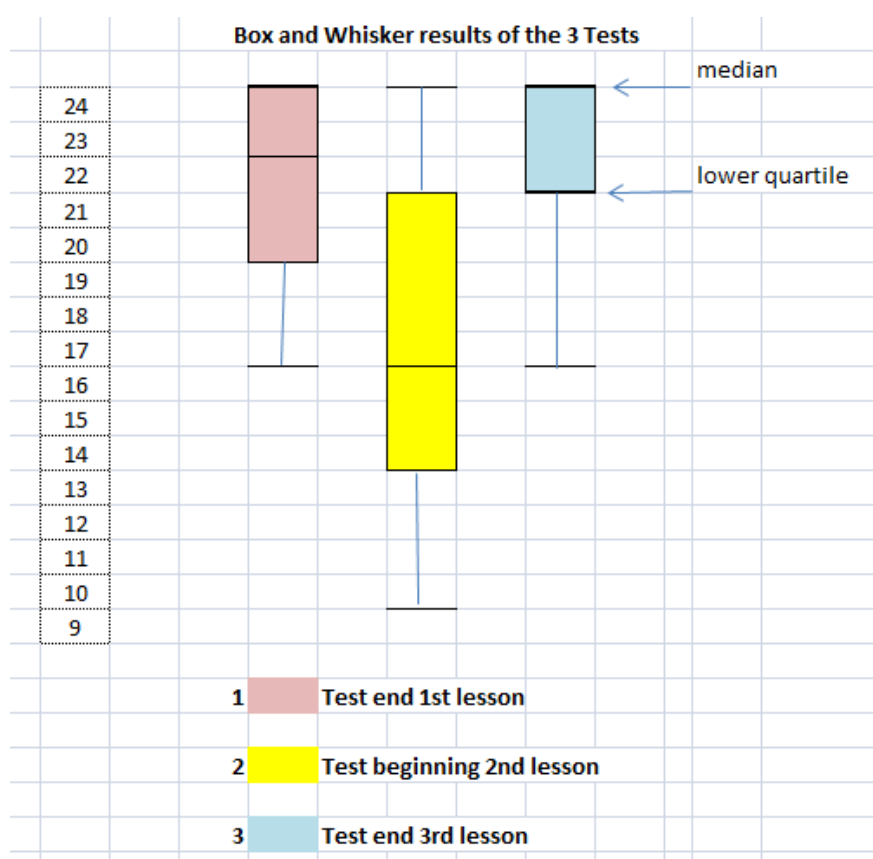
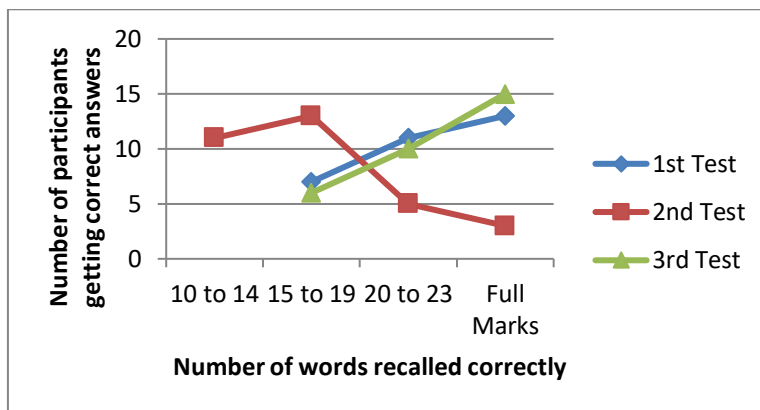


Figure 3



These are excellent results for the participant class. The second test, taken at the beginning of the second lesson, ten days later, and after an intermediary, interleaving lesson, was less successful. The retrieval rate was 75% correct answers, the mean was eighteen words retrieved correctly (out of the twenty-four) and the median was seventeen words, the upper quartile being twenty-one words and the lower, fourteen words, the highest and lowest records being the full twenty-four words and the lowest being only ten words correctly recalled (see Figures 1, 2 and 3). The spread of the results is much wider than the first test as can be seen in Figure 4.

Figure 4



Although the results are lower than the first test, they were assessed nearly two weeks after the first learning session, and are still very good when compared with the general results of this class in similar circumstances. (This comparison will be further discussed in the next chapter). The third test, carried out at the end of the second lesson of further problem-solving tasks again using pairwork, achieved excellent results. The retrieval rate was 94% correct answers, the mean, twenty-three (out of twenty-four) words retrieved correctly. The median was the top mark of twenty-four and the lower quartile, twenty-two words, with the lowest score being seventeen words (see Figures 1, 2 and 3). Thus half the class achieved between twenty-two and twenty-four correct answers (out of twenty-four).

These results correspond to the expectations of the readings and researches cited above. Creating ‘desirable difficulties’ such as the problem-solving tasks set for my participant class, slows down the long-term memory encoding process that, according to Bjork and Bjork (2011), optimises long term retention and transfer. Similarly, Cepeda *et al.*’s research (2006) demonstrated that spacing and interleaving the learning (as replicated, in the present mini-research, by the intermediary lesson on a different topic and the ten-day gap between the research lessons), show significant learning benefits. For Vygotsky, requiring participants to work together in social interaction, helps resolve cognitive difficulty and internalise information into long-term memory. Also retrieval practice, such as the testing carried out at the end of both lessons and the beginning of the second lesson, is, in itself, according to Bjork and Bjork (2011), a vital pedagogical tool for learning. The participants’ exemplary results thus correlate with the expectations of the vast majority of these previous researches, that the likelihood of obtaining better memory encoding and retrieval would be achieved by reproducing these conditions.

Conclusion

The findings of this research appear to concur with the results anticipated in the previous researches, studies and readings discussed in the literature review. However, the results must first be placed into context in order to analyse and interpret their significance. As suggested in the last chapter, it is necessary to compare the findings with habitual results based on the usual teaching methods adopted by School X. This is not straightforward, as written vocabulary retrieval, or testing, is not the prevailing teaching or assessment method in the school. It is therefore necessary to compare the study findings with retrieval results revealed by different assessment methods more widely in use in the school, such as the use of newly taught vocabulary in oral or written sentences. As this assessment method is not a systematic quantified technique producing clear factual figures, results have to be interpreted. As suggested above, it is imperative that such interpretative analysis be undertaken with great caution and circumspection.

The participant class had been taught vocabulary about school subjects and telling the time, over a period of two lessons, preceding the research project. I had formulated a personal opinion of the classes' retrieval of the new vocabulary from their oral and written work during the second lesson on this topic. I had assessed an overall average retrieval rate of around 60%, with the strongest elements of the class obtaining an 80-85% proportion of correct results, and the weakest elements obtaining 40-45%. While this remains a non-verifiable assessment, it approaches a broad and comprehensive appraisal of customary vocabulary retrieval results for this class (agreed with the class teacher).

In view of these customary results, the findings of the mini-research show a vastly improved average retrieval rate of 86% over the three tests, with half the class exceeding 90% in the final test, the median being top marks, 100%, and the weakest element obtaining 70% retrieval. A further test a month later would be necessary to verify and secure the long-term durability of the results, as Roediger *et al.* warn, in their 2011 study, that massed tests may achieve practically error-less retrieval but not provide long-term retention (Roediger *et al.*, 2011).

Taking into consideration the caveat of the non-verifiable nature of the evidence proving that the research results are greatly improved, compared with customary retrieval results for this class, it can nevertheless be reasonably argued that, with an average of 94% correct retrieval rate for the class at the end of the study, the findings of my mini-research agree with Bjork and Bjork's assertion that introducing 'desirable difficulties' enhances the learning process and strengthens the encoding of new information into long-term memory (Bjork & Bjork, 2011). Furthermore, Bjork and Bjork's insistence that retrieval (testing), rather than re-studying, makes information more retrievable again in the future is shown to be effective in this short study. It is, however, less evident to be able to confirm the added value of collaborative social interaction as this is already extensively used as a pedagogic tool in School X.

Also, my limited research was an amalgam of various ‘desirable difficulties.’ These include interleaving, thus alternating practice; cognitive conflict, when new information did not fit in with preconceived notions, such as in puzzles and problem solving, which were tackled in interaction with others (Vygotsky, 1987); spacing or distributed practice when a laps of time was created between learning during the lessons, and between the lessons; creating cognitive effort for the brain to organise random information into a meaningful sense, while modifying the context or environment of learning; and finally using testing as a tool creating memory traces through the effort of retrieval. As the research was a combination of all these added ‘difficulties’ it is therefore not possible to affirm which difficulty is the most effective in enhancing long-term memory retrieval. Therefore recommendations for future practice, based on this research, must necessarily advocate a combination of all the ‘difficulties,’ employing a methodology similar to the test.

For me, this mini-research has opened up a new teaching and learning potential, both for personal learning use and for class teaching. Like most people, I considered that re-studying, or re-reading information would fix it in my memory through greater acquaintance and familiarity with the notions embedded within the text. This research has caused me to realise that re-reading / re-studying may, in fact, be an illusion of re-learning, of very little benefit, perhaps even a waste of time. Retrieval practice develops a better memorisation and long-term retention of such notions, especially when retrieval is spaced from the original learning and interleaved with intermediary topics. When teaching, creating additional difficulties that generate cognitive conflict or necessitate re-organising random information into meaningful sense, especially when worked in social interaction, is stimulating, and also assists in encoding the information into long-term memory. In all cases, intellectual effort, as an ‘active process of interpretation,’ referencing the new knowledge onto the ideas we already know, and making sense out of chaotic data, encode the information into our long-term memory far more efficiently than traditional passive methods of ‘study’ and ‘reading’ (Bjork & Bjork, 2011).

My recommendations for future practice, therefore, based on this research, would be, when teaching new material, to set my students a combination of ‘difficulties’ ensuring intellectual effort is required to make sense of the problems posed. I would advocate the use of collaborative social interaction to work out the problems, and ensure spacing, interleaving and, in particular, information retrieval (testing) are included in the scheme of work.

Limitations and future development

My study was limited by its time frame; further retrieval practice at spaced intervals would have provided supplementary evidence as to the reliability of the results. Also, not having a control class to compare hard factual figures with my findings means that I have no hard evidence to validate or

negate my conclusions. I also did not give any feedback to my participant class about their individual results, as the test was done entirely anonymously to respect the ethical framework of the research project. Without this feedback, the participants were unable to receive the full benefit of their effort, as Roediger *et al.* conclude in their study (2011) that testing with feedback leads to much higher benefits than retrieval without feedback. Yet, even with its limitations, my mini-research produced findings which have enabled me to judge for myself the effectiveness and benefits of ‘desirable difficulties:’ problem solving, cognitive conflict, and testing, on long-term retention of new information and the implications of this effectiveness for language teaching.

The main ‘desirable difficulty’ techniques employed in this research are cognitive conflict, spacing, interleaving and testing. These techniques can be introduced into a classroom routine with little effort. Cognitive conflict and problem solving constitute already a fundamental component of language teaching when pupils are faced with new, incoherent elements which need to be organised into the familiar meaningful framework of their native tongue. ‘Deep’ questioning and imposing meaning on difficult content is also relatively straightforward to adapt to ‘normal’ teaching strategies. Interleaving, switching topics and obliging learners to reload and recode memories, would also require little effort on the part of the teacher to be put into effect. Furthermore, reviewing prior learning by intervals allows pupils to build new material onto their prior learning (Deans for Impact, 2015). As already discussed, interleaving has been found particularly efficient with mathematics learning as it gives pupils the opportunity to become used to selecting an appropriate strategy to deal with different problems, instead of simply applying the last strategy used. Interleaving and spacing, or ‘distributed practice’ necessitate programming material at intervals and returning back to topics already studied. As remarked by Bahrck *et al.*, it is a pity that current curricula rarely plan on the intermittent and repeated retrieval practice of previously acquired learning (Bahrck *et al.*, 1993) as the compelling potential offered by widely distributed retrieval practice is therefore not sufficiently exploited.

Finally, the now widely accepted value of retrieval practice as memory reinforcement, is often ignored because of the stigma attached to testing, and also, in the domain of language learning, the fear of increasing the impression that modern foreign languages are too ‘difficult’ to obtain good marks in, and should be dropped in favour of more ‘mark-rewarding’ subjects. The stigma attached to this sentiment tends to dissuade MFL departments in their pedagogical choices. Yet a strong level of evidence has been built up evidencing that quizzes and tests as a ‘delayed re-exposure’ to key information, help both its retention and its retrieval by students (Pashler *et al.*, 2007). The fact of retrieving the information from one’s long-term memory, helps fuse the data more deeply into this long-term memory, by a system of re-loading and re-coding, which then diminishes the risk of forgetting. However, the conditions under which the retrieval practice is carried out have been proved to affect the strength of the benefit. If the retrievals are massed in easy one-minute intervals, perfect recalls will not produce significant long-term benefits in retention. Roediger *et al.*’s 2011 study

concludes that the most efficient retrieval schedules are associated with spacing tests by at least a week. Testing can be completed in various ways, and not always in the classroom, for example by self-testing with set questions or flash-cards. A further benefit of testing is the improved confidence level of learners who feel better prepared for exams and more sure of their knowledge. Feedback from the tests is also important to long-term retention. Delayed feedback (answers given at the end of the test, or even later) led to better retention results, in Roediger *et al.*'s 2011 study, than immediate feedback after each question. It is thought that delayed feedback may correlate with the spacing effect (Roediger *et al.*, 2011).

As these 'desirable difficulty' techniques (cognitive conflict, spacing, interleaving and testing), are relatively simple to put into effect when teaching in school, and given the fact that I have witnessed, first-hand, the efficiency of their effects, I fully intend to use them in my own teaching practice as learning tools. The negative connotations of problem solving, cognitive conflict and testing can be offset by explaining their benefits for the pupils, not only in obtaining better retention of vocabulary, but also in their own confidence building, once the initial stress of sitting tests has passed, learners feel better prepared for further tests and examinations.

Reference List:

- Agarwal, P.K., Bain, P.M., & Chamberlain R.W. (2012) "The value of applied research: Retrieval practice improves classroom learning and recommendations from a teacher, a principal, and a scientist," *Educational Psychology Review*, 24(3), 437-448).
- Aristotle, *De Memoria et Reminiscentia*, circa 350 BC.
- Atkinson, R. C. (1975). "Mnemonotechnics in second-language learning," *American Psychologist*, 30(8), 821-828.
- Bahrack, H.P., Bahrack, L.E., Bahrack, A.S., & Pahrack P.E. (1993) "Maintenance of foreign language vocabulary and the spacing effect." *Psychological Science*, 4, 316-321.
- Bera Ethical Guidelines <https://catal0g.info/downloads/ethical-guidelines-for-educational-research-bera.pdf> (Accessed: 15/04/2019)
- Best, J. (2000). "The Role of Context on Strategic Actions in Mastermind" *Journal of General Psychology*, 52(3), 451-487.
- Bjork, R. A. (1975). "Retrieval as a memory modifier." In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*(pp. 123-144). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bjork, R.A. (1994). "Memory and metamemory considerations in the training of human beings." In J. Metcalfe and AShimamura (Eds.), *Metacognition: Knowing about Knowing* (pp. 185-205). Cambridge, MA: MIT Press.
- Bjork, E. L., & Bjork, R. A. (2011). "Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning." In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*. 56-64. New York: Worth Publishers.
- Bransford, J.D., Brown, A. L., & Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.
- Briggs, A.R.J., Coleman, M. & Morrison, M. (eds.) (2012). *Research Methods in Educational Leadership and Management* (3rd Edition). Lesage: London.
- Bucciarelli, M., Khemlani, S., & N. Johnson-Laird, P. (2008). "The psychology of moral reasoning." *Judgment and Decision Making*. 3. 121-139.
- Cepeda, N.J., Pashler, H., Vul, E., Wixted, J.T., & Rohrer, D. (2006). "Distributed practice in verbal recall tasks : A review and quantitative synthesis." *Psychological Bulletin*, 132, 354-380.
- Carpenter, S.K., Pashler, H., Wixted, J.T. & Vul, E. (2008). "The effects of tests on learning and forgetting." *Memory & Cognition* 36 (2), 438-448.
- Chandler, P., & Sweller, J. (1992). "The split-Affention Effect as a Factor in the Design of Instruction." *British Journal of Educational Psychology*, 62(2),233-246.
- Ericsson, K.A., Krampe, R.T., & Tesch-Römer, C. (1993). "The role of deliberate practice in the aquisition of expert performance." *Psychological Review*, 100(3), 363-406.

- Foskett, N., Lumby, J. And Fiddler, B. (2005). "Evolution or extinction? Reflections on the future in educational leadership and management." *Education Management Administration and Leadership*, 33(2): 245-53.
- Impact Wales (2018) Short Guide to Action Research <https://www.impact.wales/sketch-notes/> (Accessed: 15/04/2019)
- Karpicke, J.D., Butler, A.C., & Roediger, H. L. (2009). "Metacognitive strategies in student learning: Do students practice retrieval when they study on their own?" *Memory*, 17(4), 471-479.
- Kuo, T. & Hirshman, E. (1996). "Investigations of the testing effect." *American Journal of Psychology* 109, 451-464.
- Lackey, J. (1999). "Testimonial Knowledge and Transmission." *The Philosophical Quarterly*, Vol.49, No 197, 471-490.
- Lewin, K. (1946) "Action research and minority problems." *Journal of Social Issues*, 2 (4), 34-46.
- McDaniel, M. A., Hines, R. J., Waddill, P. J., & Einstein, G. O. (1994). "What makes folk tales unique: Content familiarity, causal structure, scripts, or superstructures?" *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(1), 169-184.
- Mogonea F., Popescu, A.M. (2014) "The Role of Sociocognitive Conflict in Academic-type Learning," *Procedia - Social and Behavioral Sciences*, Volume 180, 2015, Pages 865-870, ISSN 1877-0428, <https://doi.org/10.1016/j.sbspro.2015.02.228>.
- Muijs, D. & Reynolds, D. (2010). *Effective Teaching: Evidence and Practice*. Sage Publications Ltd. London. 2010
- Pashler, H., Bain, P.M., Bottge, B.A., Graesser, A., Koedinger, K., & McDaniel, M. (2007). *Organizing Instruction and Study to Improve Student Learning*. U.S. Department of Education. Washington DC: National Center for Education Research, Institute of Education Sciences.
- Peters, E. E., & Levin, J. R. (1986). "Effects of a mnemonic imagery strategy on good and poor readers' prose recall." *Reading Research Quarterly*, 21(2), 179-192.
- Piaget, J. (1975). *The development of thought : equilibration of cognitive structures*. Paris: PUF (Eng. Trans. New York: Viking Press.
- Roediger, H.L. & Butler, H.L. (2011). "The critical role of retrieval practice in long-term retention." *Trends in Cognitive Sciences*, January 2011, Vol.15, N°1
- Rohrer, D., Dedrick, R. & Stershic, S. (2014). "Interleaved Practice Improves Mathematics Learning." *Journal of Educational Psychology*. 107. 10.
- Sacco, K., Bucciarelli, M. (2008). "The role of cognitive and socio-cognitive conflict in learning to reason." *Mind & Society*. From: http://www.psych.unito.it/csc/pers/bucciarelli/pdf/2008_MindSoc.pdf
- Skoumios, M. (2009). "The Effect of Sociocognitive Conflict on Students' Dialogic Argumentation about Floating and Sinking." *International Journal of Environmental & Science Education*, 4(4), 381-399. From: <http://oaji.net/articles/987-1404720105.pdf>

- Somekh, B., & Lewin, C. (Eds;). (2011). *Theory and methods in social research*. Sage.
- Sweller, J. (1988). "Cognitive load during problem solving. Effects on learning." *Cognitive Science*, 12(2), 257-285.
- Thomas, G. (2013) *How to do Your Research Project*. 2nd ed. London:Sage.
- Toppino, T.C. & Cohen, M.S. (2009). "The testing effect and the retention interval: questions and answers." *Experimental Psychology*. 56, 252-257.
- 'Vygotsky and Second Language Acquisition' in the encyclopedia of Applied Linguistics
<https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781405198431.wbeal1272>
- Vygotsky, L. S. (1987). *Mind in society: The development of higher psychological processes*. Harvard University Press, Cambridge M.A.
- "Thinking and Speech." In R. W. Rieber, & A. S. Carton (Eds.), *The Collected Works of L. S. Vygotsky* (Vol. 1), *Problems of General Psychology* (pp. 39-285). New York: Plenum Press. (Original Work Published 1934)
- Welsh Government. February 2015. *Successful Futures: Independent Review of Curriculum and Assessment Arrangements in Wales*.