ISSN 0910-500X

英文學思潮

THOUGHT CURRENTS IN ENGLISH LITERATURE

VOLUME XCIII

2020

THE ENGLISH LITERARY SOCIETY
OF
AOYAMA GAKUIN UNIVERSITY

青山学院大学英文学会

英文學思潮

THOUGHT CURRENTS IN ENGLISH LITERATURE

VOLUME XCIII

2020

THE ENGLISH LITERARY SOCIETY
OF
AOYAMA GAKUIN UNIVERSITY

青山学院大学英文学会

Cognitive Skill Training Theories for Optimal Language Learning

Masayuki Tanabe

ABSTRACT

This paper reviews cognitive skill training theories that can contribute to language learning. The article describes five theories having strong empirical support from the experimental psychology literature, aiming at sharing those theoretical ideas with teachers and students who are not in this particular research field. The topics covered are the generation effect, the spacing effect, overlearning, prior knowledge, and task difficulty, which mainly deal with retention and transfer of skills and knowledge. For each theory, explanations of the terminology, examples of empirical support, and pedagogical implications will be given.

Keywords: skill training, retention and transfer, language learning

1. Introduction

A major challenge in education, including language teaching, is how to promote retention of knowledge provided in a classroom and its successful reproduction in practical situations. Durability and transferability of acquired skills and knowledge are fundamental aspects of learning when considering the fact that there is a certain period of disuse of what is learned until it is put into practice and that situations of learning and subsequent performance are in most cases different. Therefore, I will address major psychological theories that are able to contribute to cognitive skill training and optimal language learning, with the aim of sharing ideas with teachers and students, especially for those who are not in this area of research.

In this paper, I will describe five theories that I thought viable with solid empirical support in the experimental psychology literature. For each theory, I will give explanations about the terminology, then introduce some empirical studies that support the theory, and finally present pedagogical implications from the theory. These theories range from classic to relatively new ones. I would like readers to consider these ideas not as definite solutions but as tentatively plausible methods that are still being researched.

2. The Generation Effect

2.1. Terminology

The generation effect refers to the advanced retention of items in memory led by the learner's generating of missing parts of input during learning rather than by reading a whole thing passively. Generation effects underlie the depth-of-processing framework (Craik & Lockhart, 1972), a theory that claims that elaborate processing of input enhances superior retention. The effect was first reported by Jacoby (1978), who compared retention performance between "discovery" learning and "reception" learning. His experiments showed that word pairs learned in a generative way from cues (e.g. foot – s__e) were recalled significantly more at a posttest than simply-read pairs (e.g. foot – shoe).

2.2. Empirical support

Slamecka and Graf (1978) presented the first set of solid evidence of the generation effect. A cued-recall paradigm was used for the generation condition of their experiment. They tested the hypothesized retention advantage of self-generated items over simply-presented items. Twenty-four students at the University of Toronto studied 100 word pairs, each presented separately on an index card. For the participants in the generate condition, each card consisted of a set of stimulus words and the initial letter of a response word (e.g. rapid – f). After drawing a card from the pile, the participant generated a word according to an operative rule (e.g. synonym) that was cued orally by the experimenter (e.g. this case 'fast' can be generated). Five rules, associate, category, opposite, synonym, and rhyme were used separately for each 20 consecutive cards. For the read-only participants, an intact word pair (e.g. rapid – fast) was presented per card and the relationship of the words (e.g. synonym) was told. In both conditions, participants had to verbalize both the cue word and the response word. Prior to the study phase, they were informed that recognition of the responses would be assessed at the end. Immediately after the study phase, recognition of all target words was assessed in a three-choice format (i.e. a correct choice and two distractor choices per item) together with a 5-point confidence rating for each of them. Results showed that self-generated items were more accurately recognized than read-only items at a significant level regardless of the operating rules. Analyses of the confidence rating also revealed that the generation group chose their answers more confidently than the read-only group. This recognition superiority was further tested by comparing recognition performances with and without the notification of a subsequent recognition test at the beginning of the experiment. As a result, it was found that the difference of effect sizes from the informed–uninformed manipulation was very small with an advantage for generative processing as a whole. This led the researchers to conclude that the generation effect was so robust and that a manipulation of whether a test was notified or not was "totally inconsequential" (p. 595).

McDaniel, Howard, and Einstein (2009) proposed a study strategy for learning from educational texts, incorporating the generation effect into a practical use in education. The strategy, termed the read-recite-review technique, or the 3R in short, requires students to read texts once and recite as much of the text aloud as they remember, and then read the same text again. They contrasted the 3R, to examine its effectiveness, with two other commonly employed techniques: rereading and note-taking. In their experiments, seventy-two students at Washington University studied four short passages containing about 250 words that were taken from an EFL textbook in one of the three conditions to which they were randomly assigned. All participants were, first, notified that they would be tested about the materials later. The reread group was told to read the passages twice without any other additional activities. The note-taking group also read the passages twice, and took notes during reading to help their memory for the subsequent tests. They were not allowed to use the notes on the tests. The 3R group read the passages once, recalled the passages as well as they could, and then read the passages again. Their memory of the text was assessed in three different forms of tests: a free-recall test, a multiple-choice test, and short-answer inference questions in which could be answered by synthesizing two or more ideas from the passages. An immediate posttest asked about the first and third passages. A delayed posttest, performed a week later, was about the second and fourth passages. Analyses on immediate and delayed free-recall performances showed significantly greater recall probabilities of the 3R than those from the other study strategies. Meanwhile, the 3R group did not outperform the other study strategies on the multiple-choice test and the short answer questions. However, when more complex passages were employed, the 3R significantly benefitted on performance on a multiple-choice test, both immediate and delayed, and immediate performance on short-answer questions, as well as free-recall performance. These results suggested, according to the researchers, that generative learning was beneficial beyond improving retention. In conclusion, they proposed that students could be encouraged to utilize the 3R strategy for remembering information from class materials.

2.3. Pedagogical implications

As evidenced by the fact that Slamecka and Graf (1978) employed a cued recall paradigm to make their subjects encode stimulus word pairs (e.g. rapid – f), the generation effect largely relies on the subject's first language lexicon. Therefore, it does not seem suitable for second language learning at first glance. In fact, a second language learner does not know many words in the target language, so there is virtually no chance to come up with the correct word even when a cue was given. However, as stated in the 3R strategy (McDaniel *et al.*, 2009), the generation effect can be expected in second language learning by making effort to recall as much material as possible after initially reading it. What can be mentioned that have in common these seemingly different examples is that the generation effect could occur when input processing and recall of its relevant information happened at the same time. Therefore, when introducing new vocabulary words for example, it is important to provide ideas peripheral to the target vocabulary, so that the learner could have a chance to process both of them.

3. The Spacing Effect

3.1. Terminology

The spacing effect, researched for more than 130 years (Ebbinghaus, 1885/1913; Thorndike, 1912), refers to enhanced learning facilitated by leaving an interval between two study episodes for the same material. The interval that separates different study episodes is called an interstudy interval. Another interval called a retention interval refers to a time lag between the end of the last study episode and a subsequent test. In learning, when the same material appears twice with an interstudy interval, it is spaced practice. By contrast, when different material is inserted between the two presentations, for example, material A appears, material A disappears, material B appears, material B disappears, and then material A appears again, it is called 'distributed practice.' A typical counterpart condition to spaced practice is 'massed practice.' When the same material is presented without any interventions or rest intervals, learning is said to be massed. Superior retention performances from spaced learning that have been found in many studies were results of the comparisons with those from massed learning. A quantitative meta-analysis by Cepeda *et al.* (2006) identified 317 experiments in 184 articles in this research domain suggesting overall positive effects of spacing in learning, for which the effects in question are reasonably robust.

3.2. Empirical support

An early study on the spacing effect suggests that foreign language learning may be optimized by employing spaced practice in the long run. Barhrick, Barhrick, Barhrick, and Barhrick (1993) conducted a 9-year longitudinal investigation for four subjects who learned and relearned 300 English–French or English–German word pairs at intervals of 14, 28, and 56 days. Retention was tested one, two, three, and five years after the completion of learning. The study found that increasing the interstudy gap to 56 days improved recall performance in all retention conditions (i.e. 1 – 5 years).

Other researchers have focused on an optimal interstudy gap of a-year-long for learning. Cepeda, Vul, Rohrer, Wixted, and Pashler (2008) investigated optimal interstudy gaps for retention intervals up to one year (i.e. 7 days, 35 days, 70 days, and 350 days) based on data collected from 1,354 participants' fact learning (e.g., "What European nation consumes the most spicy Mexican food?"). They found that interstudy gaps of 1, 11, 21, and 21 days were the best for the retention intervals of 7, 35, 70, and 350 days, respectively (21-day interval benefits to both 70- and 350-day retention), suggesting that about a month-long interstudy interval would enhance long term retention from two months to a year.

Spaced practice may also be effective in abstract learning. In Kornell and Bjork's (2008) study, 120 students at University of California viewed 72 scenery paintings by 12 unfamiliar artists. Half of the participants were assigned to the massed presentation condition, and the other half were assigned to the spaced presentation condition. In the massed condition, the paintings were presented in such a way that six pictures by the same artist were given successively and then six pictures by another artist were given. In the spaced condition, the paintings by the 12 artists were intermixed. After the study phase, participants were shown 48 new paintings by the same 12 artists, and were asked to identify the name of the painter for each picture. As a result, participants in the spaced presentation condition identified the authors of those pictures significantly more accurately than participants in the massed presentation condition. This study provides a further insight into second language learning because it implies that spaced practice may also be effective for grammar learning, which involves abstraction of complex rules.

3.3. Pedagogical implications

When learning involves repeated presentation of material or a set of analogous practice stimuli, it is advisable to have an interval between training sets. This suggestion may make you focus on spaced presentations in class but spacing can be done at levels of longer periods of time than that, such as days, weeks, and months. Therefore, using a part of class time for refresher activities for learned knowledge or skills is also a viable procedure in the long run. It would be possible to say that implementation of such refresher activities should not be notified beforehand in terms of ensuring a desired period of interval. Also, when giving feedback to a learner, summary feedback after completing a set of trials is more advisable than giving trial-by-trial feedback in terms of the spacing effect.

4. Overlearning

4.1. Terminology

Overlearning refers to extra continuation of practice after reaching a set criterion at the initial stages of learning. A typical experimental manipulation for overlearning requires a learner to continue practicing beyond arriving at an errorless trial of the learning set, while a control learner stops the learning when a perfect trial was made. Research has suggested that overlearning promotes retention of information in cognitive learning as well as retention of skills in motor learning (Driskell *et al.*, 1992). Rationales for overlearning argue that transmission of knowledge from short-term memory to long-term memory is promoted by overlearning (Foriska, 1993), and that access to knowledge in long-term memory encourages effortless retrieval of the knowledge (Hall, 1989, p.328).

4.2. Empirical support

A meta-analysis study by Driskell, Willis, and Copper (1992) examined 16 overlearning studies conducted from the 1920s to the 1980s. Eleven of those studies used cognitive tasks which mainly involved learning of words, and the remaining five were based on physical tasks such as maze tracing or assembling a military gun. Typically, participants in these studies conducted learning tasks for a designated number of times, and several more trials were additionally given to treatment participants. According to analyses of effect sizes by Driskell *et al.*, effects of overlearning were moderate in size for both cognitive learning (d = .753) and physical learning (d = .443), suggesting that overlearning can enhance subsequent retention. However, it should be noted that most studies

Driskell *et al.* analyzed employed retention intervals of two weeks or less. This suggests that long-term advantages of overlearning are still not known.

More recently, Rohrer, Taylor, Pashler, Wixed, and Cepeda (2005) employed longer retention intervals of up to nine weeks for fact learning. In this study, 130 undergraduates at the University of South Florida studied 10 city-country pairs (e.g. Axim – Ghana), with either five or 20 test trials. Retention of knowledge was tested at 1, 3, and 9 weeks after learning. Results showed, in terms of statistically significant recall probabilities, that effects of overlearning lasted up to three weeks after learning. The results were consistent with the previous findings, in which overlearning was effective for short-term goals but might not be very efficient in the long run. Rohrer *et al.* mentioned that overlearning would not be efficient as previous studies had indicated, especially when it comes to remembering information for a meaningfully long term. However, because the short-term effect of overlearning was found to be reliably strong, it might be appropriate for learners who expect a short-term goal, such as students preparing for exams.

4.3. Pedagogical implications

As mentioned above, overlearning is an option that can be taken to deal with learners with an urgent need of a skill or knowledge. Therefore, it can be used more frequently in classes with a preparation purpose such as for college entrance or qualifying exams, rather than in general language classes. It also provides insights for our daily self-studying. For example, a student who is going to take a foreign language test in a few days can set up a study plan such that he or she determines the number of perfect trials to be made, and then repeats memorizing words or practicing phrases until reaching that criterion. It is a unique feature of overlearning that a target skill or knowledge can be boosted in the short term, or even from scratch, since its practice schedule anticipates making a perfect trial.

5. Prior Knowledge

5.1. Terminology

It seems plausible that a learner who knows a lot about English-related matters, for example about American culture, can acquire the English language more successfully than those who are not interested in such matters. In the field of second language acquisition, Robinson (2007) listed 'prior knowledge' as one of the influential components for successful completion of a learning task. In the skill acquisition literature, discussions of prior knowledge can be traced back to two studies that found a chess grandmaster could memorize the positions of chess pieces at a glance and reconfigure them more quickly and accurately than novice players could (Chase & Simon, 1973; de Groot, 1965).

Superior domain-specific learning performance of individuals who have a high level of knowledge in that particular domain has often been explained within a theoretical framework originally developed for text comprehension, called 'situation models' (van Dijk & Kintsch, 1983). A situation model is a mental representation of the concepts which a text describes. Also, it can be formed by many dimensions such as the integration of prior knowledge and the information provided in the text. For example, the sentence "Cathy poured water on the bonfire." makes readers generate the inference that the fire went out, based on the general knowledge that water extinguishes fire. In this way, a greater amount or quality of knowledge leads the learner to make more links to existing representations, and thus opportunities to update knowledge can occur more easily.

5.2. Empirical Support

Morris, Grunberg, Sykes, and Merrick (1981) investigated whether individuals who knew a lot about association football (i.e. soccer) could learn more game results than those who knew little about this sport. In their experiment, 38 male students at University College Swansea answered a questionnaire about football such as "Who is Manchester City's goalkeeper?" and "Who are the last English team to win the UEFA cup?", and then listened to a broadcast of game results on the radio. The broadcast lasted four minutes, including game results for four English and three Scottish divisions, 64 matches in total. Soon after learning the results, participants were asked to write down as many of the scores as they could remember on a piece of paper that had all the fixtures on it. As a result, a significant positive correlation was found between football knowledge and recall of game results. Therefore, the researchers argued that football fans could learn scores easily probably because their richer knowledge made it possible to draw immediate implications from the result, such as league positions of their clubs and others, which then could be understood as a result of the fuller cognitive activities.

Effects of prior knowledge on learning have typically been investigated by contrasting performance of high-knowledge learners with that of low-knowledge learners as described in the abovementioned Morris *et al.*'s study. Recently, in contrast, Kole and Healy (2007) investigated whether learners can use their everyday knowledge as mediators in learning unfamiliar facts. Following a situation model account, they assumed that learning of non-domain-relevant information would be facilitated by giving cues familiar to the learner, given that experts' superior learning performance was due to a number of links available between new information and existing knowledge. In their experiment, 36 students at the University of Colorado learned a set of 144 fabricated facts. Each fact was presented as a short sentence including an unfamiliar name, a verb phrase, and an object noun (e.g. "Brian Burke likes to play lacrosse."). The control variable was the degree of prior knowledge that was set at three levels. In the low knowledge condition, the participant learned fabricated facts about unfamiliar persons. In the high knowledge condition, the unfamiliar name in each sentence was replaced with a person familiar to the participant (e.g. friends or relatives). Finally, in the mediated condition, the participant first worked on paired-associate learning between names used in the unfamiliar facts and familiar individuals, and then were exposed to fabricated facts. After the learning phase, cued recall for all the fabricated facts was administered for all participants in the three conditions (e.g. Brian Burke likes to play _____). Results for the high and low knowledge groups showed a typical prior knowledge effect, with significant better recall probabilities of the high knowledge group than the low knowledge group. However, it was also found that the mediated group significantly outperformed the low knowledge group in spite of the fact that both groups studied the same unfamiliar facts. This suggests that the mediated participants might use the unfamiliar names as cues for their existing representations to activate. Kole and Healy argued that these participants were able to integrate the new facts into the representations in the long-term memory, leading to elaborate encoding of new items.

5.3. Pedagogical implications

Prior knowledge is the most important aspect for learning processes to take place given that learning involves associating a new thing with existing knowledge. The importance of prior knowledge can be articulated most clearly in the case of vocabulary learning. In theory, we have a word dictionary that is biologically available, called the 'mental lexicon.' In the mental lexicon, a word entry is associated with other entries that are conceptually, phonologically, orthographically, or syntactically related to that entry. For example, the entry of 'dog' may have links to entries such as 'animal', 'puppy', 'bark', and 'dodge'. Naturally, an object with multifaceted support will be more stable than an object with a single support. Likewise, a new word entry that has multiple links to other entries will be more stably retained than poorly linked ones in the mental lexicon. In this way, prior knowledge determines the number of scaffolds for new knowledge to survive.

6. Task Difficulty

6.1. Terminology

Task difficulty here refers to a degree of difficulty for a learner to complete a training task. Introduction of levels of difficulty into a training task is done by providing a variety of interference during training. A theoretical assumption of task difficulty is that because interference induces the learner to make erroneous responses, the learner needs to attend to input so that the material is more likely to be elaborately processed than it would be on an easy training task. The depth of encoding processing is a key to determining the quality of subsequent retention (Craik & Lockhart, 1972).

In their paper entitled *New Conceptualizations of Practice*, Schmidt and Bjork (1992) mentioned that typical training procedures that we believe effective were in fact far from optimal, provided that the long-term retention and transferability of learned knowledge were emphasized in learning. They argued that correct responses during learning could not necessarily be reliable predictors of successful acquisition, by picking up earlier studies showing that interference during training might enhance subsequent performance, such that random practice yielded better retention than blocked practice (Shea & Morgan, 1979). Minimum feedback facilitated subsequent retention performance (Schooler & Anderson, 1990), and variability in practice reduced erroneous responses on subsequent new tasks (Mannes & Kintsch, 1987). A remarkable feature of their concept is that it posed a question to the long-held belief that errorless trials during training increased acquisition. Rather, conditions that depress performance during training can maximize acquisition in the long run. In this line of research, Bjork (1994, 1999) proposed the desirable difficulty framework, arguing that learners' post-training performance becomes maximized when optimal levels of difficulty were given during training.

6.2. Empirical Support

Contextual interference is a form of task difficulty that is brought about by altering ways of item presentation in a training task. Carlson and Yaure (1990) investigated effects of contextual interference in acquisition of Boolean calculus roles. In their study, participants learned four Boolean functions (AND[a, b], OR[a, b], NAND[a, b], NOR[a, b]) in three conditions: blocked per function, randomized, and mixed (i.e. blocked for two functions and randomized for the other two). Rate of learning in each condition was measured in terms of the mean reaction time for correct responses of the rule learning. Transfer of knowledge was assessed on a separate problem-solving task which involved equation-changing calculations (e.g. Given that A = AND[1, 0] and B = OR[0, 1], then X = NOR[A, B]?). Results showed that the blocked practice condition (i.e. low interference involved) boosted rate of learning more than the random practice condition (i.e. high interference involved). However, in the transfer task, participants in the random practice condition produced correct responses significantly faster than those who practiced in the blocked practice condition. Carlson and Yaure's study suggested that interference during training promoted flexibility of learned knowledge, but another study also suggested that interference during training could enhance subsequent retention of learned knowledge. In a context similar to Carlson and Yaure, which involved learning of Boolean functions, Schneider, Healy, Ericsson, and Bourne (1995) found that random practice schedules slowed down rate of learning but yielded superior retention performance at an immediate posttest in comparison to blocked practice schedules. Also, a remarkable contrast at long-term retention was found, in which skills learned in blocked schedules decayed four weeks after the initial learning, whereas skills learned in random schedules remained relatively intact after a month.

Schneider, Healy, and Bourne (2002) manipulated levels of task difficulty with translation direction. Following the revised hierarchical model of bilingual mental lexicon (Kroll & Stewart, 1994), they assumed that foreign word learning involving translation from a first language to a second language was more difficult than that involved the reverse direction. In their experiment, 64 English-speaking students who had little knowledge about French worked on paired-associate word learning between French and English (e.g. voiture – car). Half the participants learned word pairs by answering an English word from a French cue word (e.g. voiture – ?). The other half learned word pairs by answering a French word from a English cue word (e.g. ? – car). Therefore, the former was an easy learning condition and the latter was difficult. Results showed typical Task difficulty effects. That is, rate of learning, measured with trial-by-trial accuracy during training, was better for the easy learning condition, but participants in the difficult learning condition outperformed those in the easy learning condition in rates of retention, measured at a week after learning, as well as transfer, measured with translation in the opposite direction.

6.3. Pedagogical implications

There may be many teachers who pay attention to making their talk easy to understand for students, but making everything easy may be wrong, especially in training and practice. Research suggests that training tasks become effective by introducing interference into them. Therefore, teachers should consider introducing sources of interference into classroom exercises. It is in the nature of classroom learning that what is learned in class can rarely be applied in exactly the same way to real-life situations. Therefore, ideally, transferability of knowledge needs to be considered in addition to retention. The notion of task difficulty provides an idea that a flexible skill or knowledge can be yielded by dealing with interference during an acquisition phase.

7. Conclusion

In this paper, I have reviewed five cognitive skill training strategies that could contribute to optimal language learning. A goal of language learning is to obtain skills and knowledge that can be put into practice in real-world situations. Therefore, how to gain durability and transferability in what is learned is an issue of prime concern. Discussions of optimal language learning directly relate to this very issue, and the five theories I have mentioned above can be potent solution candidates to answer them. However, as I noted at the beginning of this paper, these theories have not been fully investigated. Therefore, it is important for teachers to check whether these strategies can apply to their students effectively. Some points to consider can be students' proficiency, age, or the purpose of the class. Finally, we may need to acknowledge that ways of learning should ideally be theoretically-driven, so that we can make further investigations objectively.

References

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., and Bahrick, P. E. (1993). Maintenance of Foreign Language Vocabulary and the Spacing Effect. *Psychological Science, 4*(5), 316-321.

Bjork, R. A. (1994). Epilogue: Institutional Impediments to Effective Training. In D. Druckman & R. A. Bjork (Eds.), *Learning, Remembering, Believing: Enhancing Human Performance*. pp.295-306. DC: National Academy Press.

Bjork, R. A. (1999). Assessing Our Own Competence: Heuristics and Illusions. In D. Gopher

& A. Koriat (Eds.), *Attention and Performance XVII*. pp.435-459. UK: MIT Press.

Carlson, R. A., & Yaure, R. G. (1990). Practice Schedules and the Use of Component Skills in Problem Solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 484-496.

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*(3), 354-380.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing Effects in Learning: A Temporal Ridgeline of Optimal Retention. *Psychological Science, 19*(11), 1095-1102.

Chase, W. G., & Simon, H. A. (1973). Perception in Chess. *Cognitive Psychology, 4*, 55-81.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of Processing: A Framework for Memory Research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

de Groot, A. D. (1965). *Thought and Choice in Chess*. The Hague, Netherlands: Mouton.

Driskell, J. E., Willis, R. P., & Copper, C. (1992). Effect of Overlearning on Retention. *Journal of Applied Psychology, 77*(5), 615-622.

Ebbinghaus, H. (1885). *Memory: A Contribution to Experimental Psychology*. Ruger, H. A., & Bussenius, C. E. (Trans. in 1913). NY: Teachers College Columbia University.

Foriska, T. J. (1993). What Every Educator Should Know About Learning. *Schools in the Middle, 3*, 39-44.

Hall, J. F. (1989). *Learning and Memory* (2nd ed.). MA: Allyn & Bacon.

Jacoby, L. L. (1978). On Interpreting the Effects of Repetition: Solving a Problem Versus Remembering a Solution. *Journal of Verbal Learning and Verbal Behavior, 17*, 649-667.

Kole, J. A., & Healy, A. F. (2007). Using Prior Knowledge to Minimize Interference When Learning Large Amounts of Information. *Memory & Cognition, 35*(1), 124-137.

Kornell, N., & Bjork, R. A. (2008). Learning Concepts and Categories: Is Spacing the "Enemy of Induction"? *Psychological Science, 19*(6), 585-592.

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language, 33*(2), 149-174.

Mannes, S. M., & Kintsch, W. (1987), Knowledge Organization and Text Organization. *Cognition and Instruction, 4*, 91-115.

McDaniel, M. A., Howard, D. C., & Einstein, G. O. (2009). The Read-Recite-Review Study Strategy: Effective and Portable. *Psychological Science, 20*(4), 516-522.

Morris, P. E., Gruneberg, M. M., Sykes, R. N., & Merrick, A. (1981). Football Knowledge and the Acquisition of New Results. *British Journal of Psychology, 72*, 479-483.

Robinson, P. (2007). Criteria for Classifying and Sequencing Pedagogic Tasks. In M. P. Garcia Mayo (Ed.,) *Investigating Tasks in Formal Language Learning*. pp.7-27. UK: Multilingual Matters.

Rohrer, D., Tylor, K., Pashler, H., Wixted, J. T., & Cepeda, N. J. (2005). The Effect of Overlearning on Long-Term Retention. *Applied Cognitive Psychology, 19*, 361-374.

Schmidt, R. A., & Bjork, R. A. (1992). New Conceptualizations of Practice: Paradigms Suggest New Concepts for Training. *Psychological Science, 3*(4), 207-217.

Schooler, L. J., & Anderso, J. R. (1990). The Disruptive Potential of Immediate Feedback. In *Proceedings of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Shea, J. B., & Morgan, R. L. (1979). Contextual Interference Effects on the Acquisition, Retention, and Transfer of a Motor Skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*(2), 179-187.

Schneider, V. I., Healy, A. F., Ericsson, K. A., & Bourne, L. E. (1995). The Effects of Contextual Interference on the Acquisition and Retention of Logical Rules. In A. F. Healy & L. E. Bourne (Eds.), *Learning and Memory of Knowledge and Skills: Durability and Specificity*. pp.95-131. CA: Sage Publications.

Schneider, V. I., Healy, A. F, & Bourne, L. E. (2002). What Is Learned under Difficult Conditions Is Hard to Forget: Contextual interference Effects in Foreign Vocabulary Acquisition, Retention, and Transfer. *Journal of Memory and Language, 46*, 419-440.

Slamecka, N. J., & Graf, P. (1978). The Generation Effect: Delineation of a Phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4*(6), 592-604.

Thorndike, E. L. (1912). The curve of work. *Psychological Review, 19*, 165-194.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. NY: Academic Press.