Language Teaching Research 9,2 (2005); pp. 129–146

An investigation of the cognitive processes underlying the keyword method of foreign vocabulary learning

Amy M. Shapiro and **Dusty L. Waters** *University of Massachusetts Dartmouth*

The Keyword Method (KWM) of vocabulary learning is a mnemonic method designed to help students learn foreign vocabulary. It has been shown to be effective over several other memorization strategies. The present experiment was designed to explore the cognition underlying the effectiveness of the KWM. Specifically, both the degree of cognitive engagement with the method and visual encoding were examined as possible sources. Subjects were tested in a 2 (imagery level) × 2 (processing strategy) mixed design. Each subject was asked to memorize 30 Latin vocabulary words, divided evenly among high- and low-imagery value words. Subjects were either provided with both keywords and interactions (the Given condition) or instructions to generate their own keywords and interactions (the Self-Generated condition). Retention was tested in both immediate and delayed post-tests. Results revealed a strong effect of imagery level in both post-tests. No significant main effect of processing strategy or interactions between imagery and processing were observed in either post-test. Results indicate that the KWM is effective because it provides a meaningful visual image upon which to base memory for a new word's meaning. They suggest that there is some flexibility in how the KWM is used.

I Introduction

Students of foreign languages often rely on various strategies to memorize vocabulary words. One such method is the Keyword Method (KWM) of vocabulary learning (Atkinson, 1975; Atkinson and Raugh, 1975; Raugh and Atkinson, 1975). In using this two-step mnemonic

Address for correspondence: Amy Shapiro, Associate Professor, Department of Psychology, University of Massachusetts Dartmouth, MA, USA; e-mail: ashapiro@umassd.edu

 10.1191/1362168805lr151oa

method, a keyword is generated for the to-be-learned (TBL) word. The keyword must be a familiar word to the student and acoustically similar to some part of the TBL word (e.g., the English keyword *steel* for learning the French word *stylo*). Then, an interaction between the keyword and the meaning of the TBL word is to be visualized (e.g., imagining a large pen made of a steel girder). The interaction needs to be played out vividly in the student's mind. For example, the student should envision a famous movie star struggling to sign autographs with a huge steel girder that glimmers in the sunlight. In essence, a short movie clip is created that encompasses the two elements. This interaction may seem silly, but the more outrageous the interaction is, the more effective the method is (Solso, 1998).

A number of experiments have demonstrated the effectiveness of the KWM. One of the most common research designs has been to test the KWM against other methods of learning vocabulary. The KWM has been compared with learning words in context (e.g., immersion) or learning words with no given strategy. For example, McDaniel and Pressley (1989) found the KWM to be significantly more facilitative to learning over the context method. It has also been shown that the KWM is superior over systematic teaching (King-Sears *et al*., 1992). Additionally, this method has been shown to be an effective tool in learning many words in a short amount of time. Atkinson and Raugh (1975) found that students were able to learn 120 new words in three days (40 words a day). In sum, a good deal of research has shown that the KWM of vocabulary learning is highly effective.

Other studies have been conducted on the generalizability of the KWM for the learning disabled, garnering mixed results (Pearlman, 1990; Mastropieri *et al.,* 1990; Condus *et al.,* 1986). Mastropieri *et al*. (1990) found that primary-grade school subjects did significantly better at learning both concrete and abstract definitions of words when the KWM was used over a classical instruction condition. Even more impressively, Condus *et al*. (1986) found significant effects for the KWM over three other learning conditions (subjects were 62 12-year-olds with learning disabilities), both in an immediate recall test and one administered a full 10 weeks later. These two studies indicate that those with learning disabilities can use this method to aid in the learning of new vocabulary words. In contrast, Pearlman's (1990) study found the KWM to be too difficult for learning-disabled children in the 4th–6th grade to use; direct instruction facilitated higher recall. In Pearlman's (1990) study, however, only 16 subjects were used and they possessed differing levels of handicap. It may be that the effectiveness of this method lessens as the degree of disability increases, but it does not seem to indicate clearly that the KWM is ineffective for this population.

1 Why does the KWM work?

a Visual imagery: One reason that the KWM may be effective is because it takes advantage of the strength of visual memory. It is well documented within the cognitive literature that visual stimuli create very strong memories. In fact, a number of studies have shown that lists of concrete, easily imaged words (e.g., boat) are better remembered than more abstract words (e.g., liberty). Perhaps most convincing are studies that have shown the incredible accuracy of visual memory. For example, Shepard (1967) had subjects study pictures or sentences. Afterwards, they were each shown pairs of pictures or pairs of sentences and asked which in each pair they had previously studied. Subjects in the sentence condition displayed an 11.8% error rate, while those in the visual condition had only a 1.5% error rate on a recognition post-test. In an even more impressive demonstration, Standing (1973) had subjects study 10,000 images. In a delayed memory test, they were 83% accurate.

Given the robust nature of visual memory, it is reasonable to posit that the effectiveness of the KWM may be related to the creation of a visual image that ties the TBL word to its meaning. The KWM is not unique in its reliance on imagery, as other mnemonic devices also take advantage of robust visual memory. The method of loci, for example, takes advantage of a learner's knowledge of a familiar place to remember lists of items. To utilize the method of loci, first one must identify a familiar place (e.g., one's home) and the order of locations in which one would commonly carry out an activity in this familiar place (e.g., arriving home after work). Using this example of a very familiar place and activity, one can visualize the sequence of events that occur to learn a list of otherwise unrelated words. Specifically, upon arriving home after work every day, one may first gather that day's mail from the mail box on the front lawn, unlock the front door, walk through the hallway and then place the mail on the kitchen counter. The locations of these four sequences of events are as follows: lawn, front door, hallway and kitchen counter.

Next, the words to be remembered (e.g., that day's shopping list) are to be paired with the four locations (e.g., peanut butter with lawn, bread with front door, dog food with hallway, and pudding with kitchen counter) and then images of the pairs interacting are created (e.g., peanut butter smeared all over the front lawn and bread being thrown at the front door, etc.). When the shopping list needs to be recalled, this can be done by remembering the order in which one commonly performs the task.

Many studies show that this method can be very effective in the recall of lists. Carlson (1976) found significantly better recall when a group trained on the method of loci was compared to a control group. Another study by Roediger (1980) looked at the method of loci along with three other well-known mnemonic methods. When his results were compared to those of his control group, all four mnemonic groups recalled the 20-word list better. However, the method of loci and the peg word system were found to be better methods to use when the order of words remembered was important.

The basic idea of the peg word system is to learn a set of rhyming pairs and then peg the word to be learned to one of these pairs of words that one already knows, again creating an image between the two. For example, memorizing a list of words first involves memorizing a set of rhyming 'pegs' such as 'one is for sun, two is for zoo, three is for key'. The new TBL words are then 'pegged' onto this memorized set by creating images for each (e.g., 'apple' is pegged to 'sun' by visualizing a bright, shining apple in the sky). This system has also shown strong facilitation of recall. Morris and Read (1970) found remarkable recall using this method even after many trials in which the same words were used in differently ordered lists.

All of these methods rely on some degree of imagery, and evidence points to the imagery as a major cause of their effectiveness. In studies of visual mnemonics compared with rote learning, Desrochers (1982) and Paivio and Desrochers (1979) found that students learning French as a second language recalled more when using an imagery mnemonic. Ellis and Beaton (1993) found that the KWM produces superior recognition memory for new vocabulary. A rote strategy, however, was shown to be superior for production of new vocabulary words. They also found, however, that less imageable keywords, such as verbs, reduce the effectiveness of the KWM, while highly imageable nouns produce the best results. Overall, they report that the best results are obtained when the KWM is combined with rote learning.

Why might visual imagery be so effective in enhancing memory? An explanation that has a good deal of empirical support is Paivio's (1971) *Dual-Coding Theory*. Dual-Coding Theory proposes that memories are stored by the mind in two different coding systems, verbal and visual. While each of these codes may overlap, more emphasis is usually placed on one or the other for any given item. For example, an abstract word such as 'sensitivity' may be coded more strongly verbally, while a visual code may be dominant for concrete words such as 'mouse'. This proposal would explain Paivio and Desrochers' (1979) finding that superior memory is achieved when the KWM is paired with rote learning, as the combination of methods strengthens both the visual and verbal codes.

As discussed above, Desrochers (1982) found that the KWM was more effective when the keyword was more concrete. If the imagery created by the KWM is at the heart of the strategy, the imagery value of the TBL word itself should also alter its efficacy. This phenomenon has been shown to occur by using other methods. That is, if a noun is highly abstract or concrete, there can be a decrease or increase, respectively, in recall (Day and Bellezza, 1983; and Sheehan, 1972). For example, a noun with an abstract meaning, (e.g., justice) can be more difficult to imagine than a noun with a more concrete meaning (e.g., puppy). One can see a puppy in one's mind, but how does one see justice? One goal of the experiment reported here was to explore the degree to which imagery powers the KWM.

b Cognitive effort: Another possible cognitive basis for the success of the KWM may be that learners using the method are required to expend a greater amount of cognitive effort. A great many studies have shown that increasing cognitive effort, commonly referred to as active learning, boosts memory (Eysenck and Eysenck, 1979; Tyler *et al.*, 1979). Studies of text-based learning are full of such examples. For example, Walker *et al*. (1983) showed that sentences requiring greater effort to understand actually improved memory for the sentences' content.

McNamara *et al*. (1996) have shown that requiring greater effort of readers can boost the content of memory for entire texts. They manipulated the coherence of a text about the human heart by adding or deleting phrases that clarified the relationships between ideas. Comprehending the low-coherence text required the reader to make a number of inferences while the high-coherence text supplied its own clarifications. For instance, a sentence from the high-coherence text in one of their experiments was, 'Other drugs dissolve the lumps which break off the walls of arteries so that they do not stop the flow of blood to the heart' (1996: 41). The low-coherence version of that sentence was, 'Other drugs dissolve the lumps which break off the walls of veins and arteries' (1996: 42). McNamara *et al*. found that subjects with sufficient prior knowledge tended to score better on post-tests of meaningful learning such as inference making and problem solving when presented with the low-coherence text. In other words, subjects with sufficient prior knowledge to draw from actually did better on post-tests when there was less information in the text. The authors conclude that the effort of making sense out of the sparse text by applying what they already knew to 'fill the gaps' resulted in more robust memory.

Craik and Lockhart (1972) proposed that there are different levels of processing, ranging from shallow to deep. Shallow processing typically involves a focus on attributes other than the meaning of an item to be remembered. As such, shallow processing is typically characterized by a focus on surface features of items to be remembered such as sounds, orthography, or physical features such as the number of vertical lines in a word. Deep processing entails focusing on meaning. The richer the meaning attached to an item to be remembered, the deeper the processing and the more likely it will be recalled.

Craik and Tulving (1975) demonstrated this point in a now classic study by asking subjects to engage in shallow or deep processing of words to be remembered. Shallow processing was assigned to subjects by having them answer structural questions about each word (e.g., is the word in capital letters?). Moderate processing entailed answering phonemic questions about each word (e.g., does the word rhyme with *weight?*). A third group was asked to engage in deep processing by considering questions of a semantic nature (e.g., would the word fit the sentence 'He met a ____ in the street?'). Subjects were then tested for how many words could be recalled freely. Their results revealed that more cognitive engagement (deeper processing) took longer to accomplish. However, the deeper the information was processed, the better later recall was.

When using the KWM, the process of creating a keyword and an interaction requires a certain amount of engagement with the word and its meaning. Given the large body of evidence that expending greater effort during study results in improved memory retention, it is reasonable to posit that the KWM works because it requires learners to become actively engaged in the task of attaching meaning to the TBL words. If that is indeed the case, reducing the amount of effort required of learners to use the KWM should reduce its effectiveness. This hypothesis was tested by Hall (1988) who studied the effects of the KWM on students who had been highly trained in the use of this method. He compared given versus self-generated keywords and found only a slight effect for the Self-Generated condition (Given: mean = 15.64 and Self-Generated: mean = 16.50). However, he only tested 15 subjects in all (7 and 8 in each of the two conditions). His results, although not impressive for the Self-Generated condition, could have been affected by the small test sample. Additionally, he did not control for time spent on each word and subjects were given only four minutes in which to complete a post-test. A second goal of the present study was to recreate Hall's experiment with a larger sample size and more controlled conditions, in order to test the cognitive effort hypothesis.

II The experiment

The experiment we report here was designed to explore the cognitive processes underlying the KWM. Toward this end, we test three specific hypotheses. Our first hypothesis is that visual imagery is at the heart of the KWM. If it is, the KWM should be unable to compensate for the inherent difficulty of remembering abstract concepts. Therefore, we have hypothesized that the KWM would be insufficient for boosting memory for low-imagery definitions. After all, low-imagery words are, by definition, harder to visualize. For example, while one may visualize 'happiness' with a smiling face, conjuring the visual image during the post-test is no guarantee of being able to recall 'happiness' rather than 'smile', 'funny', or other related concepts. In contrast, visualizing 'table' with a table is likely to allow accurate recall of the correct definition. Our second hypothesis is that cognitive effort powers the KWM. In other words, the level of activity or processing required to generate a keyword and an interaction produces a more robust memory. If that is the case, reducing the amount of cognitive effort required of subjects by providing them with keywords and interactions should mitigate the effectiveness of the KWM. If both imagery and cognitive effort contribute to the KWM, we should see an interaction between these variables. Our third hypothesis, then, is that memory for new vocabulary will be best when imagery and effort are both high and worst when both are low.

We explored these hypotheses by providing subjects with Latin vocabulary words that were judged to have either high or low imagery value. The Latin words and their English meanings are provided in Appendix A. We used Latin because we had to ensure that our subjects did not know the language we chose for our study. For the purposes of our study, we did not view Latin as substantially different from modern languages and using Latin prevented the need to exclude large numbers of potential subjects. The other variable, cognitive effort, was manipulated by creating conditions in which keywords and interactions are either provided to subjects or self-generated. We explored the long-term effects of the KWM by providing a delayed post-test. The studies on the KWM have looked at results from an immediate post-test, but few extended the post-test past a few days (Wang and Thomas, 1995; McDaniel and Pressley, 1989; and McDaniel and Pressley, 1984). Immediate post-tests provide ease of memory retrieval due to their recent activation in short-term memory (Benjamin *et al.,* 1998). That ease gives way to an artificially inflated measure of true long-term memory. A more accurate measure of memory in experimental settings, then, is derived from delayed post-tests. Moreover, a delayed post-test is more ecologically valid, as the goal of language learning is long-term retention. Therefore, to delay the post-test a few days would be very informative. For these reasons, post-tests were administered immediately and after a delay of one week.

1 Participants and materials

A total of 104 students, who were enrolled in an introductory psychology course at the University of Massachusetts, participated to fulfil a requirement for that course. They were roughly evenly divided between men and women. A total of 48 subjects were used to pre-test the stimuli and 56 participated in the actual experiment.

To develop appropriate stimuli for the experiment, 103 English words were pretested for degree of imagery. This was achieved by giving a survey to the 48 pretest participants who were asked to rate each word on a scale from 1–7 for degree of imagery (higher scores indicate greater imagery value). The participants were not subjects in the actual experiment. Of the initial 103 words, 30 were chosen for use in the actual experiment. The mean for the 15 low-imagery words chosen for the study was 2.52, with a range of 2.05–2.88. The mean for the 15 high-imagery words was 6.55, with a range of 6.40–6.70. The difference between imagery sets is statistically significant ($t(28) = 58.59$, $p < 0.001$).

Keywords and interactions for all 30 of these words were created. Since Latin is the root of many English words and that of other languages commonly studied by undergraduates, all Latin words were pre-screened for similarity to words in English and other Romance languages. A professor at the University, who is fluent in Latin as well as two other Romance languages, screened the words. He is a veteran teacher, having taught foreign languages at the University of Massachusetts for over 25 years.

2 *Experimental design, procedure and apparatus*

Subjects were tested individually in a 2 (processing task) $\times$ 2 (imagery value) mixed design. All subjects were told that they would be presented with 30 vocabulary words and that they would later be asked to recall their meanings. They were instructed that in order to learn these words, their task would be to imagine vividly the meaning of each word interacting with a keyword. Depending on the condition to which they were randomly assigned, half of the subjects were provided with both keywords and interactions (the Given condition), while the others needed to generate both their own keywords and interactions (the Self-Generated condition).

SuperLab 2.0 was used to present the stimuli randomly to each subject. This program was also used later to present the post-tests. All subjects were presented with an identical set of words that consisted of an equal number of high- and low-imagery words. Subjects were given three practice trials in which they were asked to practise the activities required in their respective conditions. All subjects were informed that a computer program would display the words that they must learn and that the program would automatically display another pair of words every 16 seconds. They were able to view each pair of words for 15 seconds with a one-second rest in between words, during which time there was nothing on the screen. A number of studies have shown that presentation times longer than this can reduce the KWM's effectiveness (Calfee and Anderson, 1971; Dempster, 1987; Hall and Fuson, 1986, Johnson, 1964; Pressley, 1987). Subjects were further instructed that in order to learn these definitions, they would need to perform silently the requested cognitive activities (e.g., concentrating, visualizing, thinking, etc.). During this study phase, participants did not need to perform any other activity (e.g., writing).

Subjects in the Self-Generated condition were additionally informed that word pairs would be displayed and that they would need to generate a keyword and then an interaction according to the practised guidelines. Subjects in the Given condition were presented with word pairs, keywords and interactions and told to imagine vividly the interaction as outlined during practice sessions.

After all the words had been presented during the study phase, subjects were given a distractor task to complete. This task consisted of eight simple arithmetic problems that they were asked to solve using a pencil and paper. The purpose of the distractor task was to prevent rehearsal of the last items and to ensure that information about the last words had time to transfer from short-term memory to long-term memory. Without rehearsal, short-term memories will either decay or transfer to long-term memory within 30 seconds. The distractor task was necessary because our aim was to test the effectiveness of the KWM at enhancing the development of *long-term* memory.

Immediately after the distractor test had been administered, a post-test was given that tested the number of English word meanings that had been remembered. The post-test consisted of the 30 Latin words previously studied, with no other information such as keywords or interactions. Subjects were tested by presenting each Latin word, one at a time, on a computer screen. They were allowed to view each word as long as they needed to recall the correct English translation, although they were encouraged to work as quickly as they could. Once subjects knew the meaning, they wrote their answer on a response sheet.[1] If they did not know an answer they were told to put an 'X' in the space provided.

After writing a response, subjects pressed a button to receive the next Latin word.

We also administered a delayed post-test approximately one week after the initial test date. To reduce study or practice effects, subjects were not told in advance that they would be retested during the second session. Two different post-tests were created, each consisting of all 30 Latin words but in a different random order. These two post-tests were counterbalanced across the conditions to which participants were assigned. Half of the subjects received post-test order A first, followed by post-test order B, and the other half received post-test order B first, followed by post-test order A.

III Results of the experiment

Answer sheets were scored by comparing subjects' answers to a master sheet. Answers were considered correct if they were (1) identical to the master sheet, (2) correct but with minor spelling errors, or (3) the word's stem was correct while the ending was not identical (i.e., the subject answered 'hoped' or 'hoping' instead of 'hope'). A random sample of 20 answer sheets was scored by both investigators to obtain inter-rater reliability of the scoring procedure. A Pearson *r* analysis revealed a correlation of .987. The remainder of the answer sheets were then scored by one of the investigators.

For all analyses, a mixed two-way ANOVA was used. There were two levels of the imagery factor (high and low) and two levels of processing (Given and Self-Generated). Per cent correct was analysed for each processing and imagery condition, and for each session. The group means are reported in Table 1. There was a significant main effect of imagery

Table 1 Per cent correct by condition in the (a) immediate and (b) delayed post-tests

Processing condition	Imagery condition		
	High	Low	Total
Self-Generated	(a) 71.28	(a) 28.72	(a) 50.00
	(b) 75.15	(b) 15.54	(b) 45.35
Given	(a) 77.40	(a) 19.03	(a) 48.21
	(b) 83.50	(b) 11.88	(b) 47.69
Total	(a) 74.34	(a) 23.88	
	(b) 79.32	(b) 13.71	

on number correct in the immediate ($F(1, 54) = 135.90, p < 0.001$) and delayed ($F(1, 54) = 113.61, p < 0.001$) post-tests. This shows that the imagery value of the word being learned has a significant effect on recall. There was no significant effect of the processing task in the immediate or delayed post-tests ($F < 1$, in both cases). There was no significant interaction between these factors in the immediate post-test ($F(1, 54) = 3.34, p > 0.05$) or the delayed post-test ($F < 1$).

IV Conclusions

The effect of imagery was strong across conditions and across both testing periods. These findings indicate that the use of visual imagery is a fundamental underpinning of how the KWM works. Since the low-imagery words are harder to visualize in a way that directly relates a word's meaning to a visual image, the KWM was less effective for the low-imagery words. Indeed, subjects' performance differences between high- and low-imagery words are rather striking. After only 15 seconds of study per word, subjects were able to recall 79% of the high-imagery Latin words a full week later. In contrast, they were only able to recall about 14% of the low-imagery words after a week (less than one of the 15 low-imagery words they studied). This result supports our first hypothesis that the creation of a visual image, which serves as a cue to a word's meaning, is an underpinning of the KWM.

The experiment also tested whether cognitive effort may underlie the KWM's effectiveness. The analyses reveal no support for that contention. There were no statistically significant differences between processing conditions on the immediate or delayed post-tests. In spite of our large subject sample and experimental controls, this finding replicates the findings of Hall (1988), discussed previously, who also found no difference in post-test performance between given and self-generated subject groups.

Our final hypothesis was that there would be an interaction between the imagery value of the TBL words and processing strategy. We predicted that subjects would perform best when asked to self-generate interactions for high-imagery words (high effort and strong imagery) and worst when given interactions for low-imagery words (low effort and low imagery). Analyses revealed no support for this hypothesis. Combined with the nonsignificant main effect of cognitive effort, this result provides further evidence against the role of cognitive effort as an underpinning of the KWM. Instead, the very strong significant effect of imagery value points to the creation of visual imagery as the key element in the KWM. That is, it does not seem to matter how the image is created (i.e., given or self-generated), as long as the image can be easily connected with the correct definition of the TBL word. Since the connection between visual imagery and a TBL word's meaning is the crucial element, we propose that any method of creating such meaningful associations will be helpful in augmenting retention of new vocabulary.

There is some real world evidence for the validity of this conclusion that stems from a different domain of language learning, early literacy. A problem facing beginning readers (typically in kindergarten and first grade) is learning to associate the sounds of language with the letters of the alphabet. An early literacy programme used by many American school systems is the *Telian* method. The Telian method works by providing children with visual images that tie the sounds of each letter to its shape through a short, highly visual story, just as the KWM uses visual images to tie a TBL word to its meaning.

For example, the lower-case letters b and p are physically and phonemically similar. Both are bilabial sounds (both lips are pressed together to make the sound). The difference between b and p is that the b sound is voiced (the vocal cords vibrate) while the p sound is not voiced. They are physically different mainly by virtue of where the curve is drawn on the vertical line (high or low). Children are taught to recognize each as separate letters and sounds through stories and pictures. Both are called 'lip poppers' (which is easier for children than the formal, linguistic term bilabial). The letter b is associated with a baby in the Telian system, as the children are told that 'The noisy, little baby turns her voice on when she puts her lips together and puffs out the air.' The letter p is associated with the parent by telling the children that 'The tall, quiet mother keeps her voice off when she puts her lips together and puffs out the air.' Along with these visual images, the children are shown pictures of the letter b with a baby's face drawn in the curve, while a mother's face drawn in the p.

We have seen children in the early stages of reading actively calling upon the images and stories as they attempt to sound out letters, even when not prompted to do so. The system is very effective and is currently in use by a large number of school systems across the USA.

The success of the programme is relevant to this discussion because it works in spite of the fact that the children are *provided* with the images. In other words, the association between a letter and its sound through a visual image is what makes the Telian system work, rather than increasing cognitive effort by having children engage in *generating* an association.

In summary, prior research on the KWM has made clear the fact that this method helps students learn new vocabulary words. The contribution of the present research effort is to show that the KWM's power does not arise from the increased cognitive effort of generating keywords and interactions. Rather, it appears to stem from use of visual imagery that ties a new word to its meaning. The fact that the KWM does not require a high level of cognitive effort to be effective is good news for students, because it means that the method can be made easier for them to use without jeopardizing learning outcomes. This finding translates to greater flexibility with respect to how the KWM is used by instructors and students of foreign languages. As such, there may be some opportunity for creative use of the KWM both in and outside of the classroom. For example, an instructor may also choose to provide keywords and interactions that are designed to make students laugh and enjoy the class, while still enjoying the proven benefits of the KWM. For example, a 'word of the day' can be provided each day that will motivate students to come to class. One can imagine students on the way to class wondering, 'What crazy scene will he have for us today?' The method can also be turned into a game in which the entire class participates. For example, each student can be given one or more words and assigned the task of coming in with keywords and interactions to present to the class. In this way, class members enjoy themselves, get to know their classmates, and help each other with vocabulary. The exercise can be turned into a contest for the most outrageous or memorable images. Such games can be very motivating to students and the present research suggests that they are pedagogically sound.

Note

1 Subjects were asked to write their responses rather than type them because there was great variability in typing skill among our population sample. We were concerned that the added task of 'hunting and pecking' would be distracting to a few subjects and might interfere with their performance on the task at hand.

V References

Atkinson, R.C. 1975: Mnemotechnics in second-language learning. *American Psychologist* 30: 821–28.

Atkinson, R.C. and **Raugh, M.R.** 1975: An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory* 104: 126–33.

Benjamin, A.S., Bjork, R.A. and **Schwartz, B.L.** 1998: The mismeasure of memory: when retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General* 127(1): 55–68.

Bjork, R.A. 1979: Information-processing analysis of college teaching. *Educational Psychologist* 14: 15–23.

Calfee, R.C. and **Anderson, R.** 1971: Presentation rate effects in paired-associate learning. *Journal of Experimental Psychology* 88(2): 239–45.

Canellopoulou, M. and **Richardson, J.T.** 1998: The role of executive function in imagery mnemonics: evidence from multiple sclerosis. *Neuropsychologia* 36(11): 1181–88.

Carlson, R.F., Kincaid, J.P., Lance, S. and **Hodgson, T.** 1976: Spontaneous use of mnemonics and grade point average. *Journal of Psychology* 92(1): 117–22.

Condus, M.M., Marshall, K.-J. and **Miller, S.R.** 1986: Effects of the keyword mnemonic strategy on vocabulary acquisition and maintenance by learning disabled children. *Journal of Learning Disabilities* 19(10): 609–13.

Craik, F.I.M. and **Lockhart, R.S.** 1972: Levels of processing: a frame-work for memory research. *Journal of Verbal Learning and Verbal Behavior* 11: 671–84.

Craik, F.I.M. and **Tulving, E.** 1975: Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 104: 268–94.

Day, J.C. and **Bellezza, F.S.** 1983: The relation between visual imagery mediators and recall. *Memory and Cognition* 11(3): 251–57.

Dempster, F.N. 1987: Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology* 79(2): 162–70.

Desrochers, A. 1982. Imagery elaboration and the recall of French article-noun pairs. *Canadian Journal of Psychology* 36(4): 641–54.

Ellis, N.C. and **Beaton, A.** 1993: Factors affecting foreign language vocabulary: imagery keyword mediators and phonological short-term memory. *Quarterly Journal of Experimental Psychology* 46A(3): 533–58.

Eysenck, M. and **Eysenck, M.** 1979: Processing depth, elaboration of encoding, memory stores, and expended processing capacity. *Journal of Experimental Psychology: Human Learning and Memory* 5: 472–84.

Hall, J.W. 1988: On the utility of the keyword mnemonic for vocabulary learning. *Journal of Educational Psychology* 80(4): 554–62.

Hall, J.W. and **Fuson, K.C.** 1986: Presentation rates in experiments on mnemonics: a methodological note. *Journal of Educational Psychology* 78(3): 233–34.

Hintzman, D.L. 1974: Theoretical implications of the spacing effect. In Solso, R.L. editor, *Theories in cognitive psychology: The Loyola Symposium*, Potomac, MD: Erlbaum, 77–99.

Johnson, N.F. 1964: The functional relationship between amount learned and frequency vs. rate vs. total time of exposure of verbal materials. *Journal of Verbal Learning and Verbal Behavior* 3: 502–04.

King-Sears, M.E., **Mercer, C.D.** and **Sindelar, P.T.** 1992: Toward independence with keyword mnemonics: a strategy for science vocabulary instruction. *Remedial and Special Education* 13(5): 22–23.

Levin, J.R., **Levin, M.E.**, **Glasman, L.D.** and **Nordwall, M.B.** 1992: Mnemonic vocabulary instruction: additional effectiveness evidence. *Contemporary Educational Psychology* 17(2): 156–74.

McDaniel M.A. and **Pressley, M.** 1984: Putting the keyword method in context. *Journal of Educational Psychology* 76(4): 598–609.

—— 1989: Keyword and context instruction of new vocabulary: effects on text comprehension and memory. *Journal of Educational Psychology* 81(2): 204–13.

McNamara, D.S., **Kintsch, E.**, **Songer, N.** and **Kintsch, W.** 1996: Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction* 14: 1–43.

Manalo, E. 1999: Spontaneous mnemonic use in simulated foreign word learning. *Psychologia: An International Journal of Psychology in the Orient* 42(3): 160–69.

Mastropieri, M.A., **Scruggs, T.E.** and **Fulk, B.M.** 1990: Teaching abstract vocabulary with the keyword method: effects on recall and comprehension. *Journal of Learning Disabilities* 23(2): 92–96.

Morris, P.E. and **Reid, R.L.** 1970: The repeated use of mnemonic imagery. *Psychonomic Science* 20(6): 337–38.

Paivio, A. 1971: *Imagery and verbal processes*. New York: Holt, Rinehart and Winston.

Paivio, A. and **Desrochers, A.** 1979: Effects of an imagery mnemonic on second language recall and comprehension. *Canadian Journal of Psychology* 33(1): 17–28.

Pearlman, I. 1990: Effectiveness of keyword versus direct instruction on vocabulary acquisition by primary-grade handicapped learners. *Bulletin of the Psychonomic Society* 28(1): 14–16.

Pressley, M. 1987: Are keyword method effects limited to slow presentation rates?: an empirically based reply to Hall and Fuson (1986). *Journal of Educational Psychology* 79(3): 333–35.

Raugh, M.R. and **Atkinson, R.C.** 1975: A mnemonic method for learning a second-language vocabulary. *Journal of Educational Psychology* 67: 1–16.

Roediger, H.L. 1980: The effectiveness of four mnemonics in ordering recall. *Journal of Experimental Psychology: Human Learning and Memory* 6(5): 558–67.

Sheehan, P.W. 1972: Role of imagery in incidental learning: replication and extension of an effect. *Journal of Experimental Psychology* 95(1): 226–28.

Shepard, R. 1967: Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior* 6: 156–63.

Solso, R.L. 1998: *Cognitive psychology,* fifth edition. Needham Heights, MA: Allyn and Bacon.

Solso, R.L. and **Biersdorff, K.K.** 1975: Recall under conditions of cumulative cues. *Journal of General Psychology* 93: 233–46.

Standing, L. 1973: Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology* 25: 207–22.

Stubin, E.J., Heimer, W.I. and **Tatz, S.J.** 1970: Total time and presentation time in paired-associate learning. *Journal of Experimental Psychology* 84: 308–10.

Tyler, S., Hertel, P., McCallum, M. and **Ellis, H.** 1979: Cognitive effort and memory. *Journal of Experimental Psychology: Human Learning and Memory* 5: 607–17.

Walker, N., Jones, J. and **Mar, H.** 1983: Encoding processes and the recall of text. *Memory and Cognition* 11: 275–82.

Wang, A.Y. and **Thomas, M.H.** 1995: Effect of keywords on long-term retention: help or hindrance? *Journal of Educational Psychology* 87(3): 468–75.

Waugh, N.C. and **Norman, D.A.** 1965: Primary memory. *Psychological Review* 72: 89–104.

Appendix A: The Latin words used in the study, paired with their English translations

High-imagery words	Low-imagery words
Arm = Bracchuim	Bribe = Donis
Box = Arca	Curse = Defixionis
Clock = Pallium	Deserter = Perfuga
Couch = Lectus	Flattery = Blanditiae
Cup = Poculum	Foolishness = Stultita
Dish = Lancis	Harm = Laedo
Door = Ianua	Hope = Spes
Ear = Auris	Injustice = Iniuria
Face = Oris	Mercy = Venia
Knee = Genu	Misfortune = Casus
Money = Opes	Mumble = Susurro
Parrot = Psittacus	Order = Lubeo
River = Flumen	Plot = Coniurationis
Shoe = Calceus	Revenge = Ultonis
Torch = Facis	Trickery = Dolus