Long-term retention under conditions of intentional learning and the keyword mnemonic

ALVIN Y. WANG, MARGARET H. THOMAS,
CAROLYN M. INZANA, and LAURIE J. PRIMICERIO
University of Central Florida, Orlando, Florida

This study was designed to assess long-term retention following either keyword learning or rote rehearsal under conditions of intentional learning. A 2 condition (keyword vs. rote rehearsal) × time (immediate vs. delay) completely randomized design was used, with 24 Tagalog language nouns serving as the learning material. All subjects had uniformly high levels of retention on the immediate test of cued recall. However, on the delayed test of cued recall (2 days later), keyword learners had forgotten almost twice the number of items as had subjects who had rehearsed the same material by rote. Long-term forgetting was also greater for the keyword condition than for rote rehearsal when subjects were assessed by a test of associative matching. This pattern of results replicates earlier research that has revealed a detrimental long-term effect of the keyword mnemonic under conditions of incidental learning.

Considerable evidence exists for the beneficial effect of the keyword mnemonic on acquisition and immediate recall. Although originally developed by Richard Atkinson (1975) as a study strategy for second language vocabulary learning, its usefulness has been documented across a wide range of learning tasks. The acquisition of science concepts, art history, the names of state capitals, and factual information presented in a prose format are but a few examples of its broad applicability (Carney, Levin, & Morrison, 1988; Levin, Shriberg, Miller, McCormick, & Levin, 1980; McCormick, & Levin, 1987; Pressley, Levin, & Delaney, 1982; Rosenheck, Levin, & Levin, 1989).

Whereas the immediate advantage of using keywords is well documented, the possible long-term benefit of this mnemonic technique has not been established. In fact, some evidence suggests that the keyword mnemonic may be associated with a relatively steep decline in retention if recall is assessed after a delay (McDaniel, Pressley, & Dunay, 1987; Wang, Thomas, & Ouellette, 1992). Wang et al. (1992) conducted four experiments to compare long-term retention after keyword learning with that after rote rehearsal. We were somewhat surprised to find that after 1 week, *greater forgetting occurred under use of the keyword mnemonic* than under rote rehearsal.

In the aforementioned experiments, condition (keyword vs. rote rehearsal) and time (immediate vs. delay) were factorially manipulated as between-subjects variables. Despite variations in the study materials (French or Tagalog language nouns), and levels of immediate recall (equated or not equated across learning conditions), keyword learners typically forgot more than twice the number of

vocabulary items than did subjects who rote rehearsed the same material. In each of these experiments, a highly significant condition × time interaction was obtained because keyword learners were less likely to retain as much after a delay as rote learners. Moreover, the poor long-term performance of the keyword learners was not due to their forgetting of the keywords. For instance, after a 1-week delay, keyword recall was 90%, as opposed to a meager 25% for English translations (Wang et al., 1992, Experiment 2). Nor was this steep decrement due to an inability of subjects to successfully apply the keyword mnemonic to the language stimuli. In all cases, immediate recall by keyword learners was very high (over 90%), and, unless rote learners were given more acquisition trials, it significantly exceeded that of the rote group. Also, debriefing-questionnaire data indicated that subjects used the suggested learning strategies correctly.

In all four of Wang et al.'s (1992) experiments, an incidental learning paradigm was used to obtain a relatively pure measure of forgetting. It was reasoned that if subjects were not anticipating a recall test, they would be less likely to practice the language stimuli during the retention interval. This was especially critical for the subjects assigned to the delayed recall condition if one assumes that different learning strategies vary in the ease with which they can be reinstated during the delay interval. Debriefing questionnaires indicated that the cover story used to promote incidental learning was successful; typically, less than 20% of all subjects indicated that they had thought about the study materials during the retention period.

Our aim in the present study was to replicate our earlier findings under conditions of intentional learning. Two related issues prompted a need for this attempt. First, it was important to establish whether the long-term forgetting associated with the keyword mnemonic would also be

Correspondence should be addressed to A. Y. Wang, Department of Psychology, University of Central Florida, Orlando, FL 32816-1390.

found under conditions that more closely reflected the demands of everyday learning and memory tasks (e.g., intentional learning). If intentional keyword learning should still produce steep declines in delayed recall, it could be concluded that keyword-based memories are especially prone to forgetting regardless of the intent to learn. Such a replication would lend greater "ecological validity" to our earlier findings. The second related issue concerned the possibility that mnemonic learning strategies might benefit more from practice than nonmnemonic techniques (Wood, 1967). To the extent that an intentional learning set should promote spontaneous practice, improvement in the long-term retention of the keyword condition relative to rote rehearsal would be expected. In this scenario, the particular type of condition × time interaction reported in our earlier experiments would not be replicated in the present study.

METHOD

Subjects

The subjects were students recruited from a general psychology class. They received extra class credit for their participation and were randomly assigned to one of the four experimental conditions ($N = 89$). The data from 2 subjects who were familiar with the Tagalog language were excluded from the data analysis.

Design and Materials

A 2 (condition) × 2 (time) completely randomized design was used. The subjects learned the language stimuli by either the keyword method or rote rehearsal and were tested either immediately or after a 2-day delay. The 24 language stimuli were concrete nouns from Tagalog (the national language of the Philippines). For the keyword learners, each language stimulus consisted of a Tagalog-keyword-English translation (e.g., MANOK-man-CHICKEN). The subjects in the rote rehearsal condition studied the same Tagalog-English word pairs, but without the keywords (e.g., MANOK-CHICKEN).

Procedure

The keyword subjects received instructions and examples for using this mnemonic technique. Two familiarization trials were then given, to strengthen the phonological association between each Tagalog word and its keyword. During the familiarization trials, each Tagalog-keyword pair (e.g., MANOK-man) was presented visually and pronounced out loud by the experimenter. The presentation rate was 4 sec for each familiarization stimulus. Following familiarization, the language stimuli (Tagalog-keyword-English) were presented across two acquisition trials at a 10-sec rate. The keyword subjects were encouraged to form vivid and interactive images combining each keyword and English translation. Different random orders were generated for each of the acquisition and familiarization trials, so that sequence effects were minimized.

The subjects in the rote rehearsal condition received 10 acquisition trials in which the Tagalog-English word pairs were presented at a 3-sec rate. Each of these language stimuli was presented visually and pronounced out loud by the experimenter. Again, different random orders

were generated for each acquisition trial to reduce sequence effects. The relatively fast presentation rate was used to inhibit the spontaneous occurrence of learning strategies other than rote rehearsal (Bugelski, 1970; Hall, 1988). Pilot studies suggested that the acquisition procedures used in both learning conditions would lead to comparable levels of immediate recall.

After the acquisition phase, all subjects were given a written 3-min general knowledge test (e.g., "Was Frank Sinatra born in Hoboken or Sicily?"), which served as a distractor task. The immediate test of retention was administered after completion of this task.

Retention was assessed via cued recall followed by an associative matching test. For keyword learners, each cued recall item consisted of a Tagalog word followed by two blank lines. The subjects were asked to write down the appropriate keyword and English translation on the lines next to each Tagalog word. The subjects in the rote rehearsal condition were presented with a cued recall test consisting of a Tagalog word followed by a single blank line. They were instructed to write down the English translation on the appropriate line next to each Tagalog word. The test of cued recall was timed for 5 min. Next, all subjects received the test of associative matching. In this test, on sheets of paper, a list of the Tagalog words and another list of their English translations were shown in different random orders. The subjects were asked to "match up" each Tagalog word with its English translation and were given 5 min to complete this task.

Because an intentional learning paradigm was used, all subjects were informed at the outset that a test for recall would be given at the completion of the experiment. Delay subjects were reminded to return in 2 days so that their recall of the language stimuli could be tested. This intentional learning set proved successful: the results of a debriefing questionnaire indicated that virtually all subjects expected to be tested for recall. Table 1 shows the percentage of subjects who expected a test of recall and who had thought about the study materials during the retention period. Overall, these percentages were substantially higher than those in our earlier studies with an incidental learning paradigm (Wang et al., 1992).

RESULTS

Figure 1 shows mean cued recall as a function of learning condition and retention interval. Visual inspection suggests that, as expected, the immediate recall levels for both learning conditions were highly comparable. However, after 2 days, the performance decrement for the keyword condition was much greater than for rote rehearsal. A 2 (condition) × 2 (time) analysis of variance (ANOVA) revealed main effects of condition [$F(1,85) = 4.94, MS_e = 26.24, p < .03$] and time [$F(1,85) = 36.92, p < .001$], as well as a significant two-way interaction [$F(1,85) = 3.83, p < .06$]. Just as in our earlier studies (Wang et al., 1992), about twice as many items were forgotten under conditions of keyword learning as under rote rehearsal.

The group means for the test of associative matching were: keyword-immediate ($M = 23.16 \pm 2.12$), keyword-delay ($M = 16.07 \pm 6.61$), rote-immediate ($M = 20.80 \pm 3.35$), and rote-delay ($M = 19.25 \pm 5.63$). A 2 (condition) × 2 (time) ANOVA yielded a main effect of time [$F(1,85) = 15.79, MS_e = 25.42, p < .001$], as well as a significant condition × time interaction [$F(1,85) = 6.49, p < .02$]. A comparison of these group means indicates that after the delay, the subjects in the rote rehearsal condition forgot only 1.55 items, whereas the keyword learners forgot 7.09 items.

Finally, an analysis of keyword recall indicated that there was a slight decrement after a delay ($M = 20.27$

Table 1
Percentage of Subjects Who Expected the Test of Recall and Who Had Thought About the Study Material During the Retention Interval

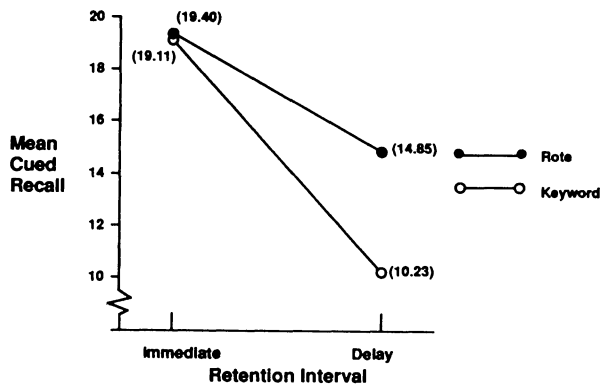|	Condition			
Debriefing Question	Rote Immediate	Rote Delay	Keyword Immediate	Keyword Delay
Expected test?	90.0	96.0	93.5	92.3
Thought about material?	80.0	50.0	65.7	26.3

Figure 1. Mean cued recall as a function of learning condition and retention interval.

± 4.45) in comparison with immediate keyword recall (M = 22.32 ± 2.10) [$F(1,47)$ = 3.51, MS_e = 13.91, p < .08]. However, the fact that delayed keyword recall remained quite high (84.5%) suggests that the forgetting of keywords was not responsible for the poor performance of keyword learners on the delayed tests of cued recall and associative matching.

DISCUSSION

The overall pattern of findings replicates those from earlier experiments in which we compared forgetting under conditions of keyword learning and rote rehearsal. It seems that regardless of whether learning is intentional or incidental, keyword-based memories are especially prone to forgetting after a delay. Just as in a previous study (Wang et al., 1992), about twice as many items were forgotten by keyword learners as by rote learners when retention was assessed by cued recall. Nor can it be claimed that the fragility of keyword-based memories is due to the forgetting of the keywords. In the present study, the keyword learners recalled almost 85% of the keywords after a 2-day delay, whereas their recall of English translations was only about 40%.

How might the meager long-term retention of keyword learners be explained? One promising account is based on an interpretive forgetting model developed by Hasher and Johnson (1975). This model distinguishes between two types of retrieval cues: contextual, temporal cues that are transitory in nature and core, semantic cues that are more retrievable over time. We believe that imagery instructions promote the formation of contextual and temporal encodings that may serve as cues or tags for keyword-based memories (Wang & Thomas, 1992). Because many of these cues are still available immediately after learning, the acquisition and recall performance of keyword learners will benefit in the short-term. However, to the extent that contextual and temporal cues become less retrievable over time, the long-term recall of keyword-based images becomes increasingly problematic. Indeed, the delayed associative matching performance of keyword learners suggests that many keyword-based memories are simply unavailable after an extended period of time.

Interestingly, rote learners were much more likely to claim that they had thought about the study material over the delay than were keyword

learners (50.5% vs. 26.3%, respectively). These findings suggest that keyword learners do not benefit as much from spontaneous practice during the delay as do subjects who learn by rate. Further research is needed to determine whether learners are more likely to spontaneously practice certain learning strategies as opposed to others. After all, the effectiveness of even the most sophisticated learning strategy will be obscured if learners prefer not to practice it over time.

The pedagogical implication of the present findings is that teachers may not assume that keyword mastery of vocabulary material in the classroom will automatically confer long-term benefits beyond rote rehearsal. Perhaps the keyword mnemonic is a specific case of Schmidt and Bjork's (1992) general observation that conditions that maximize acquisition performance do not necessarily produce effective performance after a delay. If this is the case, educators should not rely strictly on pedagogical devices that promote rapid learning in the classroom, but should attempt to foster and evaluate performance after extended periods of retention.

REFERENCES

ATKINSON, R. C. (1975). Mnemotechnics in second language learning. *American Psychologist*, **30**, 821-828.

BUGELSKI, B. R. (1970). Words and things and images. *American Psychologist*, **25**, 1002-1012.

CARNEY, R. N., LEVIN, J. R., & MORRISON, C. R. (1988). Mnemonic learning of artists and their paintings. *American Educational Research Journal*, **25**, 107-125.

HALL, J. W. (1988). On the utility of the keyword mnemonic for vocabulary learning. *Journal of Educational Psychology*, **80**, 554-562.

HASHER, L., & JOHNSON, M. K. (1975). Interpretative factors in forgetting. *Journal of Experimental Psychology: Human Learning & Memory*, **1**, 567-575.

LEVIN, J. R., SHRIBERG, L. K., MILLER, G. E., McCORMICK, C. B., & LEVIN, B. B. (1980). The keyword method in the classroom: How to remember the states and their capitols. *Journal of Educational Psychology*, **71**, 583-594.

McCORMICK, C. R., & LEVIN, J. R. (1987). Mnemonic prose-learning strategies. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications*. New York: Springer-Verlag.

McDANIEL, M. A., PRESSLEY, M., & DUNAY, P. K. (1987). Long-term retention of vocabulary after keyword and context learning. *Journal of Educational Psychology*, **79**, 87-89.

PRESSLEY, M., LEVIN, J. R., & DELANEY, H. D. (1982). The mnemonic keyword method. *Review of Educational Research*, **52**, 61-91.

ROSENHECK, M. B., LEVIN, M. E., & LEVIN, J. (1989). Learning botany concepts mnemonically: Seeing the forest *and* the trees. *Journal of Educational Psychology*, **81**, 196-203.

SCHMIDT, R. A., & BJORK, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, **3**, 207-217.

WANG, A. Y., & THOMAS, M. H. (1992). The effect of imagery-based mnemonics on the long-term retention of Chinese characters. *Language Learning*, **42**, 359-376.

WANG, A. Y., THOMAS, M. H., & OUELLETTE, J. A. (1992). Keyword mnemonic and retention of second-language vocabulary words. *Journal of Educational Psychology*, **84**, 520-528.

WOOD, G. (1967). Mnemonic systems in recall. *Journal of Experimental Psychology*, **58**, 1-27.