Influences of the encoding instructions on retrieval processes in recall and recognition memory

HIDETSUGU TAJIKA

Department of Psychology, Faculty of Education, Aichi University of Education, Kariya, Aichi 448

The purpose of this study was to explore the influences of the encoding instructions on retrieval processes in recall and recognition memory. A set of materials used consisted of low intralist similarity items in Experiment 1 (60 undergraduates), and high intralist similarity items in Experiment 2 (60 undergraduates). In both experiments, one of three retrieval tests was given to the subjects: recall, recognition, or cued recall test. Four types of encoding instruction were crossed with three types of retrieval test. The results showed that subjects performed at the greatest level when the encoding instruction and the retrieval test were compatible, and that cued recall performance was similar to recognition performance regardless of intralist similarity. It was suggested that the results were explained better in relation to the episodic theory and retrieval attributes in recall and recognition—rather than the two-stage theory of recall.

Key words: recall (learning), recognition (learning), retrieval, processes, attributes.

Many recent investigations of retrieval processes in recall and recognition memory have been based on either of two different theoretical positions. One position, the two-stage theory of recall (e.g., Anderson & Bower, 1972; Kintsch, 1970), proposes that recall and recognition memory involve basically different memory processes. Recall consists of two successive or overlapping stages, one "a search process stage" and the other "a decision process stage". In recognition, on the other hand, the search process is excluded.

The other type of the theoretical position of retrieval processes in recall and recognition memory is referred to as the episodic theory (e.g., Tulving, 1976; Tulving & Thomson, 1973; Watkins & Tulving, 1975). The episodic theory assumes that a to-be-remembered item is encoded with respect to the context in which it is stored, and that recall and recognition performances are determined by that context. As a result, retrieval processes in recall and recognition are fundamentally different in the utilization of stored information. The difference is only in the nature of the retrieval cue information that is presented at the time of retrieval. In recall, the retrieval cue information is provided with such cues as serial position and extralist items. On the other hand, the retrieval cue information in recognition is provided by a literal copy of the to-be-remembered items, which is prepared by the experimenter.

Previous results of retrieval processes in recall and recognition have been interpreted in terms of either of these two theories. For instance, Estes and DaPolito (1967) found that the subjects given instructions for intentional learning recalled more target items than those given incidental learning instructions. On the other hand, there was no difference between these two instructional groups in recognition. It suggests the possible interaction between encoding instruction and...
retrieval test. The two-stage theory of recall can relevantly explain these results.

However, there is another instance of interaction between encoding instruction and retrieval test. Tversky (1973, 1974) showed that the changes in the context between the encoding instruction and the retrieval test produced an impairment in recognition performance. This type of interaction cannot be directly explained by the two-stage theory of recall. These results can be rather handled by the episodic theory. Recall and recognition performances are determined by the compatibility of the encoding instructions and the retrieval test. When the retrieval test is compatible with the encoding instructions, both correct recall and recognition are expected to be higher than when not.

The present study was designed to extend the findings of Tversky (1973, 1974). Generally, previous research has examined the effects of test-expectancy by manipulating the encoding instructions and the retrieval test (e.g., Balota & Neely, 1980; Carey & Lockhart, 1973; Tversky, 1973, 1974). For example, Balota and Neely (1980) investigated whether the subjects who expect a recall test encode the information differently as compared with those who expect a recognition test. They presented the subjects with many items which were manipulated levels of word frequency. The results of Balota and Neely (1980) showed that the subjects who expected a recall test recalled and recognized more high frequency words than those who expected a recognition test. Moreover, the other results indicated that the subjects who expected a recall test did not recall and recognize more low frequency words than those who expected a recognition test. On the basis of the interaction between test-expectancy instruction and level of word frequency, Balota and Neely (1980) interpreted the above results within the framework of Anderson and Bower's theory (1972, 1974). It is well known that high frequency words are more likely to have multiple meanings than low frequency words (e.g., Reder, Anderson, & Bjork, 1974). It might well be said that high frequency words generally activate more memory nodes than low frequency words. Therefore, the subjects who expect a recall test may encode high frequency words more elaborately, and thus are likely to remember the to-be-remembered items correctly regardless of incompatibility of the test-expectancy instructions and the retrieval test.

The interaction between test-expectancy instruction and retrieval test has not been obtained in Balota and Neely. These findings are dissimilar to those reported by the other research (e.g., Carey & Lockhart, 1973; Tversky, 1973, 1974). These contradictory results point to the necessity for more extensive examination of an interaction between test-expectancy instruction and retrieval test.

To extend the findings of Tversky (1973, 1974), two kinds of additional variable were manipulated in our studies as follows. First, stimulus similarity of the items was manipulated by constructing low similarity items in Experiment 1, whereas high similarity items were used in Experiment 2. Second, in addition to a recall test and a recognition test used in the previous studies, a cued recall test was employed in which one or two initial letters of items served as specific cues. The experimental paradigm is similar to Balota and Neely's experiments (1980) in which each practice trial is preceded by the critical test trial.

One major merit of manipulating stimulus similarity of the items is that it is possible to extend the generality of the results of Tversky (1973, 1974). If the results found by Tversky are obtained in the present experiments regardless of stimulus similarity of the items, it is difficult for the two-stage theory of recall to account for them. According to the two-stage theory of recall, the test-expectancy instruction
Influences of the encoding instructions

variable affects only the search stage of recall. As a result, it is expected that performance of the subjects who receive a recall test would be a dissimilar pattern to that of the subjects who receive a recognition test, regardless of whether the intralist similarity of the items is high or low.

Experiment 1

Method

Subjects. Sixty undergraduates from Aichi University of Education participated in this experiment, who were paid for their service. Five subjects were randomly assigned to each of the 12 experimental and control groups. None of the subjects had previously served as subjects in a study that involved similar types of memory experiment.

Design and materials. The experiments was 4×3 between-subjects design in which the type of test-expectancy instruction was crossed with the type of retrieval test. The former consisted of free recall test-expectancy instruction, recognition test-expectancy instruction, cued recall test-expectancy instruction, and no test-expectancy instruction. The latter consisted of a free recall test, a recognition test, and a cued recall test. Subjects in each of test-expectancy instructions received three successive practice lists. The fourth study-test trial was a critical trial. In each practice list, subjects were tested their test-expectancy instruction condition. Subjects in the no test-expectancy instruction condition were tested only by the fourth test trial. Each subject in the no test-expectancy instruction condition was not instructed what kind of retrieval test was given.

Three different 30-syllable lists were selected from the Umemoto, Morikawa, and Ibuki’s norm (1955). They were two-letter syllables with 0 to 4% association value. Another 90 syllables were used as distractors in recognition tests. In a cued recall test, each initial letter was used as cues. These three 30-syllable lists were constructed to serve as practice lists. The critical test list was constructed by 30 nouns from the Koyanagi, Ishikawa, Ohkubo, and Ishii’s norm (1960). Another 30 nouns were selected as distractors in a recognition test. All nouns had a familiarity value rating in the range 3.50–4.99.

Procedure. Subjects were run individually. They were randomly assigned to one of four test-expectancy instruction conditions which differed in the type of encoding instruction. Subjects were asked to memorize a list of the 30 syllables presented successively at a 2-s rate by a projector-tachistoscope with electronic shutter. When having learned 30 items, subjects received the retrieval test appropriate to each test-expectancy instruction. They had three minutes to complete one of each retrieval test. After having studied three practice lists, subjects were given the critical test list constructed by 30 three-letter nouns with low intralist similarity. Another group of 20 subjects rated intralist similarity of the items used in this experiment on a scale from 1 (very dissimilar) to 5 (very similar). The mean rating scale of this stimulus item set was 2.60. These critical test items were presented successively at a 3-s rate. After having studied the critical test list, subjects were tested by one of three types of test, which was referred to as the first retrieval test. Subjects were allowed three minutes to complete each retrieval test. After the first retrieval test, subjects were given the interpolated task such as an arithmetic task, followed by the second retrieval test. Subjects tested by either recognition or cued recall received a free recall test, whereas those tested by free recall received a recognition test. It took three minutes for each subject to solve the interpolated task. The second retrieval test was administered within three minutes as well as that of the first retrieval test.

Results

In analyzing the results of the present
Table 1
Mean proportion of correct recall, recognition, and cued recall in Experiments 1 and 2 as a function of encoding instructions

<table>
<thead>
<tr>
<th>Encoding instruction</th>
<th>Retrieval test</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recall</td>
<td>Recognition</td>
<td>Cued recall</td>
<td></td>
</tr>
<tr>
<td>Experiment 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recall</td>
<td>.65</td>
<td>.71</td>
<td>.53</td>
<td></td>
</tr>
<tr>
<td>Recognition</td>
<td>.49</td>
<td>.86</td>
<td>.63</td>
<td></td>
</tr>
<tr>
<td>Cued recall</td>
<td>.45</td>
<td>.95</td>
<td>.71</td>
<td></td>
</tr>
<tr>
<td>No instruction</td>
<td>.48</td>
<td>.84</td>
<td>.63</td>
<td></td>
</tr>
<tr>
<td>Experiment 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recall</td>
<td>.26</td>
<td>.34</td>
<td>.25</td>
<td></td>
</tr>
<tr>
<td>Recognition</td>
<td>.34</td>
<td>.51</td>
<td>.23</td>
<td></td>
</tr>
<tr>
<td>Cued recall</td>
<td>.27</td>
<td>.53</td>
<td>.27</td>
<td></td>
</tr>
<tr>
<td>No instruction</td>
<td>.29</td>
<td>.40</td>
<td>.21</td>
<td></td>
</tr>
</tbody>
</table>

In a 4(test-expectancy instruction) × 3 (retrieval test) analysis of variance, there was a significant main effect of retrieval test, F(2,48)=47.73. The amount of correct response was greater under a recognition test than under each of the other retrieval tests. Subjects tested in cued recall performed better than those in recall. Moreover, the interaction between test-expectancy instruction and retrieval test was also significant, F(6,48)=4.53. There was no significant main effect of test-expectancy instruction, F<1.50.

As the interaction between test-expectancy instruction and retrieval test was significant, a 2(test-expectancy instruction) × 2(retrieval test) analysis of variance was carried out under three sets of condition. These sets were composed of each of the two variables compatible with the test-expectancy instruction and the retrieval test. As a 2(Recall, Recognition) × 2(Recall, Recognition) analysis of variance revealed a significant main effect of retrieval test, F(1, 16)=15.48. Subjects under a recognition test produced more items than those under a recall test. There was also a significant interaction between test-expectancy instruction and retrieval test, F(1, 16)=8.00. However, there was no significant main effect of test-expectancy instruction, F<1.00. An analysis of the interaction indicated that there was no significant effect of retrieval test under the recall test-expectancy instruction condition, but there was a significant effect of retrieval test under the recognition test-expectancy instruction condition, F(1,16)=22.83. Subjects tested in recognition produced more items than those in recall. Moreover, there were marginally significant effects of test-expectancy instruction under recall and recognition test conditions, F(1, 16)=3.99 in recall, and F(1, 16)=3.99 in recognition, both ps<.10. Subjects who expected a recall test produced more items under a recall test than those who expected a recognition test. On the other hand, subjects who expected a recognition test pro-
Influences of the encoding instructions

Table 2
Mean proportions of response classes in Experiments 1 and 2

<table>
<thead>
<tr>
<th>Encoding instruction</th>
<th>Rc (First)-Rn (Second)</th>
<th>Rn (First)-Rc (Second)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Rc, Rn)</td>
<td>(Rc, Rn)</td>
</tr>
<tr>
<td><strong>Experiment 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recall</td>
<td>.64</td>
<td>.01</td>
</tr>
<tr>
<td>Recognition</td>
<td>.48</td>
<td>.01</td>
</tr>
<tr>
<td>No instruction</td>
<td>.48</td>
<td>.00</td>
</tr>
<tr>
<td><strong>Experiment 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recall</td>
<td>.16</td>
<td>.10</td>
</tr>
<tr>
<td>Recognition</td>
<td>.28</td>
<td>.06</td>
</tr>
<tr>
<td>No instruction</td>
<td>.27</td>
<td>.02</td>
</tr>
</tbody>
</table>

Note. (Rc, Rn) or (Rn, Rc)=target items both recalled and recognized, (Rc, Rn) or (Rn, Rc)=target
items recalled but not recognized, (Rc, Rn) or (Rn, Rc)=target items recognized but not re-
called, (Rc, Rn) or (Rn, Rc)=target items neither recalled nor recognized.

produced more items under a recognition test
than those who expected a recall test. A
2(Recognition, Cued Recall) x 2(Recognition, Cued Recall) analysis of variance
showed that a main effect of test-expect-
ancy instruction was significant, \(F(1, 16) = 5.36\). It was observed that subjects who
expected a cued recall test was higher in
performance than those who expected a
recognition test. There was also a signifi-
cant main effect of retrieval test, \(F(1, 16) = 48.55\). Performance of subjects who
received a recognition test was superior to
that of subjects who received a cued recall
test. There was no significant interaction
between test-expectancy instruction and retrieval test under this set of condi-
tions. In a 2(Recall, Cued Recall) x 2(Recall, Cued Recall) analysis of variance, a
main effect of retrieval test was marginally
significant, \(F(1, 16) = 4.01, p < .10\). Per-
f ormance of subjects who received a cued
recall test tended to be superior to that of
subjects who received a recall test. There
was also a significant test-expectancy in-
struction x retrieval test interaction, \(F(1, 16) = 24.61\). An analysis of this interaction
showed that there were significant effects
of retrieval test under both recall and cued
recall test-expectancy instruction condi-
tions, \(F(1, 16) = 24.24\) and \(F(1, 16) = 13.64\),
respectively. The mean proportions of
correct responses for subjects who re-
ceived no test-expectancy instruction were
.48 in recall, .84 in recognition, and .63 in
cued recall. Recall performance of each
subject under the no test-expectancy in-
struction condition was equal to that under
the recognition or cued recall test-expect-
ancy instruction condition. Moreover, the
pattern of recognition performance for
subjects under the no test-expectancy in-
struction condition was similar to that for
subjects under the recognition test-expect-
ancy instruction condition.

The relation between the first retrieval test and
the second retrieval test. Table 2 shows the
proportions of correctly produced items
based on the reduction method of Tulv-
ing and Watkins (1975). In the left-hand
of Table 2, the first retrieval test was re-
call and the second retrieval test was rec-
ognition, and vice versa in the right-hand
of Table 2. The following types of re-
response are included in Table 2 which rep-
resents the proportions of target items; (i)
target items both recalled and recognized
(Rc, Rn), (ii) target items recalled but not
recognized (Rc, Rn), (iii) target items re-
cognized but not recalled (Rc, Rn), (iv)
target items neither recalled nor recognized (Rc, Rn). Similarly, the following types of response were composed of on the right-hand of Table 2: (Rn, Rc), (Rn, Rc), (Rn, Rc), and (Rn, Rc). Table 2 showed that the proportion of (Rc, Rn) was more than 20% regardless of each test-expectancy instruction condition, whereas that of (Rn, Rc) was about null.

Experiment 2

Method

Subjects. Subjects for this experiment were 60 undergraduates from Aichi University of Education, who were paid for their participation. None had been tested in Experiment 1.

Design and materials. Experimental design was identical to that of Experiment 1. The only difference was that the first retrieval test list consisted of 30 high similarity items. Each item was constructed with three letters which were selected and permuted from a group of six letters, and consisted of much familiar one like that in Experiment 1. The mean rating scale of intralist similarity of the items in this set was 3.82.

Procedure. The procedure followed was identical with that in Experiment 1.

Results

The first retrieval test. The mean proportions of correctly recalled, recognized, or cued recalled items for each test-expectancy instruction are given the lower half of Table 1. When the test-expectancy instruction was compatible with the retrieval test, the mean proportions were .26 under the recall test-expectancy instruction condition, .51 under the recognition test-expectancy instruction condition, and .27 under the cued recall test-expectancy instruction condition. A 4(test-expectancy instruction)×3(retrieval test) analysis of variance revealed a significant main effect of test-expectancy instruction, F(3, 48) = 4.44. The test-expectancy instructions of recognition and cued recall produced higher responses than those of the other two tests. There was also a significant main effect of retrieval test, F(2, 48) = 44.69. Subjects who received recognition were superior to those who received the other retrieval tests. Furthermore, an interaction between test-expectancy instruction and retrieval test was significant, F(6, 48) = 2.80.

As the interaction was significant, a two-way analysis of variance was conducted on test-expectancy instruction and retrieval test in a manner similar to that in Experiment 1. A 2(Recall, Recognition)×2(Recall, Recognition) analysis of variance revealed both significant main effects of test-expectancy instruction and of retrieval test, F(1, 16) = 15.26 for the former variable and F(1, 16) = 15.26 for the latter. There was no significant interaction. In a 2(Recognition, Cued Recall)×2(Recognition, Cued Recall) analysis of variance, there was a significant main effect of retrieval test, F(1, 16) = 38.07. Performance on recognition was superior to that on cued recall. No other significant effects were found in these results. A 2(Recall, Cued Recall)×2(Recall, Cued Recall) analysis of variance showed that there were no significant main effects and interaction. The mean proportions of the no test-expectancy instruction condition was .29 in recall, .40 in recognition, and .21 in cued recall. In recognition, subjects who were tested under the no test-expectancy instruction condition performed no better than those under the recognition test-expectancy instruction condition.

The relation between the first retrieval test and the second retrieval test. The lower half of Table 2 showed the results of response classes of recall and recognition as well as those in Experiment 1. While the mean proportions of correct responses decreased, the proportions of (Rc, Rn) and (Rn, Rc) were both higher than those in Experiment 1. It was clear that the proportion of (Rc, Rn) was .10 for subjects who re-
Influences of the encoding instructions

Discussion

The present experiments basically showed two aspects of the results. First, subjects performed at the greatest level when the test-expectancy instructions and the retrieval test were compatible, regardless of whether stimulus similarity of the items was high or low. Second, performance under the cued recall test condition demonstrated a similar pattern to that under the recognition test condition.

The first aspect of the results indicated that each target item was encoded with respect to the test-expectancy instructions in which it was stored, and that recall and recognition performances were determined by that context. For example, subjects who expected a recall test performed highest under a recall test in Experiment 1, and subjects who expected a recognition test did so under a recognition test in Experiment 2.

The results of Experiment 1 have been observed by Tversky (1973, 1974). It is possible to extend the results of Tversky (1973, 1974) on the basis of those of Experiment 2. These results can not be explained by the two-stage theory of recall (e.g., Anderson & Bower, 1972; Kintsch, 1970). According to this theory, subjects who expect a recognition test should have judged simply the newness of the successive test items regardless of the test-expectancy instructions. As a result, recognition performance of subjects who expect a recognition test might be similar to that of subjects who expect a recall test. On the other hand, performance of subjects who receive a recall test is a dissimilar pattern to that of subjects who receive a recognition test. These suppositions about recall and recognition performances were not confirmed by these two experiments which manipulated intralist similarity of the items.

The first aspect of the results can be accounted for by the episodic theory (e.g., Tulving, 1976; Tulving & Thomson, 1973; Watkins & Tulving, 1975). As Tulving and Thomson (1973) have argued, subjects encode target items with the test-expectancy instructions which are given at encoding. The compatibility between encoding and retrieval conditions produces the most successful recollection of target items.

However, there is one problem with the above argument. One of the results in each experiment has shown that subjects can not produce the most improvement in performance, even if the encoding instructions and the retrieval test are compatible. For example, subjects who received a recognition test in Experiment 1 performed at the greatest level under the recognition or the cued recall test-expectancy instruction. This result is consistent with that of Tajika (1980) that this type of cued recall represents basically similar retrieval attributes of stored items to recognition. Moreover, subjects who received a recall test in Experiment 2 produced no improvement in performance under each of the four test-expectancy instructions. A possible interpretation of this result is that high intralist similarity of the items reduces performance of cued recall at retrieval.

The second findings of the present experiments are consistent with the proposal of Tajika (1980) stated earlier. Cued recall used in the present experiments is assumed to consist of the same retrieval attributes as recognition, so that the pattern of performance under recall is similar to that under recognition.

Analyses of the reduction method indicated that subjects who expected recognition in the first retrieval test recognized more items in the second retrieval test. On the other hand, subjects who expected recall in the first retrieval test showed no improvement in performance in the second retrieval test. These findings support the view that retrieval attributes of recall are
different from those of recognition (Tajika, 1980; Underwood, Boruch, & Malmi, 1978).

References


(Received June 3, 1982; accepted July 14, 1984)