Levels-of-processing effects on Chinese character completion: The importance of lexical processing and test cue

YUH-SHOW LEE
National Chung-Cheng University, Chiayi, Taiwan, Republic of China

YEE-MEE CHEUNG
Chinese University of Hong Kong, Shatin, New Territories, Hong Kong

and

LEE H. WURM
Wayne State University, Detroit, Michigan

The levels of processing (LoP) effect was studied in Chinese character completion tasks. The same stem cue, either graphemically or semantically related to the target character, was used to generate either a perceptual or a conceptual test. Participants received either direct or indirect instructions, so that the same stem cue also produced either an explicit or an implicit test of memory. This allowed us to examine simultaneously the roles of perceptual versus conceptual test and retrieval intentionality in the LoP effect. No LoP effects were found when the memory test was implicit. LoP effects were also not found for stems and characters related graphemically (perceptually), but were found for stems and characters related semantically (conceptually) when the memory test was explicit. The results are discussed in terms of the importance of lexical processing for character completion tests and the importance of reinstatement of the study process during retrieval for explicit memory tests.

Numerous cognitive psychologists have devoted themselves to studying the unconscious influences of learning and memory (for reviews, see Reber, 1989; Schacter, 1987). In particular, the development of indirect or implicit measures of memory has made a significant contribution to our understanding of the diversity and complexity of human memory. Using traditional explicit measures of memory such as recognition, recall, and cued recall, participants are asked to recollect consciously something about a prior episode. In these tests, instructions during the testing refer to a particular spatial or temporal context in the participants' past experiences. In other words, the task instructions make explicit reference to an episode in the participants' personal history (Schacter, 1987). By contrast, in indirect, or implicit, memory tests, participants are told to process and respond to stimuli without reference to past experiences. The influence of memory is indicated by the improvement in performance from the previously studied episode. Because participants respond to stimuli without reference to the past, the improvement in performance is considered to reflect unconscious influence of memory.

A number of variables have been found to affect performance on explicit memory tests but not on implicit memory tests. For example, studies have shown preserved memory in memory-disordered patients on implicit memory tests but impaired memory on explicit memory tests. Other variables such as encoding manipulations during study, the generation effect, the modality effect, and memory development have also been found to dissociate implicit memory tests from explicit memory tests (for reviews, see Richardson-Klavehn & Bjork, 1988; Roediger, 1990; Roediger & McDermott, 1993).

One of the most important dissociations between explicit and implicit tests consists in the levels of processing (LoP) effect (Richardson-Klavehn & Bjork, 1988). LoP effects refer to better memory performance for material studied under a semantic (deep) condition than under a physical or nonsemantic (shallow) condition. Deeper processing leads to better performance in the subsequent memory test (Craik & Lockhart, 1972). Explicit memory tests typically show a large LoP effect, whereas null or attenuated effects of LoP are found on implicit memory tests, such as repetition priming (Brown & Mitchell, 1994; Graf & Mandler, 1984; Graf, Mandler, & Haden, 1982), lexical decision (Monsell, 1985), picture naming latency, and perceptual identification of pictures (Carroll, Byrne, & Kirsner, 1985).

We thank Richard Marsh, Alan Brown, Angaj Thapar, and an anonymous reviewer for helpful comments on the earlier version of this article, and Oi-Man Kwok and Hi-Yi Tsui for their assistance in testing participants. Correspondence should be addressed to Y.-S. Lee, Department of Psychology, National Chung-Cheng University, Chiayi 621, Taiwan, Republic of China (re-mail: psyys@ecunix.ecu.edu.tw).

Copyright 2000 Psychonomic Society, Inc.
Nonetheless, some studies have also revealed a differential effect of LoP on conceptual and perceptual implicit memory tests. In conceptual implicit tests, testing cues and target items overlap at the conceptual level, whereas in perceptual indirect tests, the overlap is at the perceptual level. The test cue in the perceptual implicit tests is a perceptually degraded form of the target that participants have to decipher. On the other hand, in conceptual implicit tests, there is no perceptual resemblance between the test cue and target item. A significant LoP effect has been found in conceptual implicit tasks in some studies, even though no effect was found in perceptual implicit tests (e.g., Blaxton, 1989; Roediger, 1990; Roediger, Weldon, & Challis, 1989). Furthermore, even for the perceptual implicit tests, the effects of conceptual processing were not all clear (see Brown & Mitchell, 1994; Thapar & Greene, 1994).

Several explanations have been proposed in order to account for these findings. Three hypotheses have been proposed to explain the effects of conceptual processing on perceptual implicit tests (Richardson-Klavehn & Gardiner, 1998). First, some studies suggest that perceptual implicit tests can also be affected by prior conceptual manipulations. This is the conceptual processing hypothesis, which suggests that implicit perceptual tests are sensitive not only to perceptual processing but also to conceptual processing ( Hirshman, Snodgrass, Mindes, & Feenan, 1990; Srivivas & Roediger, 1990; Weldon & Massaro, 1996; Weldon, Roediger, Beitel, & Johnston, 1995). Second, the voluntary contamination hypothesis proposes that the LoP and generation effects found on implicit perceptual tests indicate a contamination of voluntary retrieval during testing (Jacoby, Toth, & Yonelinas, 1993; Toth, Reingold, & Jacoby, 1994). For example, in Bowers and Schafer (1990), no LoP effect occurred in the word stem completion task when participants were unaware of the relation between the study and the testing phase. Once participants were aware of the relation, they used the word stem as a recall cue and an LoP effect was found.

Finally, the LoP effect can also be accounted for by the lexical processing hypothesis (Challis & Brodbeck, 1992; Roediger & McDermott, 1993). This account suggests that lexical processing is necessary for priming to occur (Hayman & Jacoby, 1989; Weldon, 1991). In the LoP manipulation, some forms of shallow processing do not require participants to treat studied words as lexical units. As a result, shallow processing produces no priming or less priming than does deep processing. According to this view, the LoP effect found on implicit tasks is due to a lack of the lexical e. leading in the shallow processing condition, and not due to an enhancement of conceptual processing in the deep processing condition.

With respect to conceptual implicit tests, some studies suggest an LoP effect parallel to that found on explicit memory tests. Certain factors influence conscious and unconscious retrieval in the same way (e.g., Craik & Mandler, 1984; Jacoby & Dallas, 1981; MacLeod, 1989), which amounts to the conceptual processing hypothesis. The voluntary contamination hypothesis says that there is conscious contamination in conceptual implicit tests, and that the effects of conceptual processing in these tests thus are due to the conscious contamination.

As can be seen from the foregoing, both the conceptual processing and the voluntary contamination hypotheses can explain the LoP effect found on conceptual and perceptual implicit tests. The lexical processing hypothesis, on the other hand, draws support mainly from implicit perceptual tests (Hayman & Jacoby, 1989; Weldon, 1991). Regardless of which hypothesis receives the stronger support, the current evidence suggests that two variables are crucial for determining when the LoP effect will appear. The first variable is retrieval intentionality, which characterizes the main difference between explicit and implicit tests. Explicit tests are more likely to show an LoP effect than are implicit tests. The second variable is test type (i.e., conceptual vs. perceptual tests). Tests that involve conceptual processing are more likely to show an LoP effect than are tests that involve perceptual processing.

Evidence bearing on these issues also comes from explicit memory tests. In some studies, explicit tests have shown no effect of LoP when the test has been focused on lexical or phonemic properties of words rather than on their meanings (Fisher & Craik, 1977; McDaniel, Friedman, & Bourne, 1978; Morris, Bransford, & Franks, 1977).

In the present study, we used the Chinese character completion task to investigate LoP effects. This task allowed us, first, to orthogonally manipulate retrieval intentionality and conceptual versus perceptual tests; and second, to use exactly the same test cues in all experimental conditions. This provided us with a chance to obtain a clearer result regarding the influences of the above two variables on the effects of LoP.

In the Chinese language, which is a logographic system, the symbols (characters) typically represent the basic unit of meaning. In alphabetic writing systems such as English, the symbols (letters) correspond more to the basic unit of pronunciation. Chinese characters consist of a spatial arrangement of strokes. These strokes combine to form radicals, which further combine to form characters. Different characters vary in complexity and construction, in terms of the number of strokes and the arrangement of components. Each Chinese character consists of a compact, square pattern of strokes (see the study characters in Table 1).

Chinese radicals carry either the phonological or the semantic information of the character in which they occur, but this information is not always transparent. The most common structure of a Chinese character comprises two radicals, arranged horizontally so that one is on the left and one on the right. The left (semantic) radical usually provides the semantic cueing function, indicating the general semantic category to which the overall character belongs. The right (phonetic) radical provides a guide to the pronunciation of the character in which it occurs. About 85%-90% of Chinese characters are phonetic com-
<table>
<thead>
<tr>
<th>Sample Materials for Stem Completion Test: Experiments 1-2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Semantic</strong></td>
</tr>
<tr>
<td>Study character</td>
</tr>
<tr>
<td>指 海 婦</td>
</tr>
<tr>
<td>(finger) (sea) (woman)</td>
</tr>
<tr>
<td>Stem, as a semantic cue</td>
</tr>
<tr>
<td>才 水 女</td>
</tr>
<tr>
<td>(hand) (water) (female)</td>
</tr>
<tr>
<td><strong>Graphemic</strong></td>
</tr>
<tr>
<td>Study character</td>
</tr>
<tr>
<td>捷 滅 嫌</td>
</tr>
<tr>
<td>(fast) (subtract) (dislike)</td>
</tr>
<tr>
<td>Stem, as a semantic cue</td>
</tr>
<tr>
<td>才 水 女</td>
</tr>
<tr>
<td>(hand) (water) (female)</td>
</tr>
</tbody>
</table>

Note - Within the same column, the semantic cue and graphemic cue are exactly the same left radical. The difference is in the study characters. In the experiments, the meanings of characters or stems were shown in parentheses.

pounds (Zhou, 1978), with one radical implying the meaning and the other one indicating the pronunciation. In about 90% of the phonetic compound characters, the semantic radical is on the left and the phonetic radical is on the right (Flores d’Areias, Saito, & Kawakami, 1995). All of the stimuli in the present study were of this type.

The absolute size of a character as it appears in newspaper is approximately the width of three English letters as they appear on this printed page. The interactive activation model of McClelland and Rumelhart (1981; Rumelhart & McClelland, 1982) assumes that information covering a region in space at least large enough to contain a four-letter English word is processed simultaneously. This suggests that the whole Chinese character is processed simultaneously in normal reading, and indeed it is impossible for a native speaker to process only half of the character (i.e., only the right or the left radical). The whole character must be processed simultaneously.

In a typical word completion task, participants study a list of words under various conditions and are then given cues consisting of letters and spaces for missing letters. They are required to fill in the missing letters. Performance on this task depends more on one’s retrieving perceptual information than on retrieving meaning (Jacoby, 1983a; Roediger et al., 1989).

In the Chinese character completion task, on the other hand, the stem given to participants during testing can be manipulated so that it is perceptually, phonologically, or semantically related to the target character. If the stem is the left (semantic) radical, the semantic relationship between the left radical and whole character can be manipulated. When the meaning of the left radical and the whole character are related, the left radical provides a semantic cue for the target character (see the examples in the top half of Table 1). When the meaning of the left radical and the whole character are unrelated, the left radical provides a graphemic cue (see examples in the bottom half of Table 1). A character completion task with semantic cues, which overlap with the character at the conceptual level, can be considered a conceptual memory test. On the other hand, a task with graphemic (perceptual) cues, which overlap with the character only at the perceptual level, is a perceptual test. It should be kept in mind that for Chinese character completion tasks, exactly the same test cues can be used in both perceptual and conceptual tasks (again see the examples in Table 1).

Two experiments are reported in the present study. In Experiment 1, we examined whether a priming effect can be found in the Chinese character completion task when participants are provided with semantic cues. Three levels of processing during study were manipulated. In Experiment 2, we manipulated retrieval intentionality and semantic versus graphemic cues orthogonally. We used a different set of LoP manipulations to ensure the generality of our conclusions.

**EXPERIMENT 1**

In Experiment 1, we used the Chinese character completion task to examine the LoP effect. Intentional versus incidental retrieval and three levels of processing (Physical, Read, and Generate) were manipulated. Given the findings in previous studies, we predicted that LoP effects would be found in the intentional retrieval testing condition. We expected that the explicit stem cue recall test would show a typical LoP effect, with performance best in the generate condition, worst in the physical condition, and intermediate in the read condition. If the conceptual processing hypothesis or the voluntary contamination hypothesis was correct, the same pattern would also be found in the implicit stem completion task. On the other hand, the lexical processing hypothesis predicted that there should be no difference between the generate and read conditions, and that the performance in physical condition should be worse than that in both.

**Method**

**Participants and Design.** Seventy-two Chinese University of Hong Kong undergraduates participated in the present experiment. They were all paid for participating and were native Chinese (Cantonese) speakers who also spoke English.

The experiment was a mixed factorial design with one between-subjects factor, test instruction (stem completion vs. stem cue recall test), and one within-subjects factor, study condition (physical vs. read vs. generate). Forty-eight participants were randomly assigned.
to a test instruction group, and the other 24 participants were assigned to the baseline condition. Participants in the baseline condition did only the stem completion test without studying the words.

**Materials.** The materials were 30 horizontally structured Chinese characters consisting of left and right radicals. None of them shared the same left radicals, which indicate the meaning of the character, and none of them shared right radicals, which indicate the pronunciation of the character. The testing materials for the stem completion test were characters with the right radical removed. There were several possible ways to complete each stem (see the examples in Table 1).

These 30 characters were divided into three groups which were used in physical, read, and generate study conditions. The three groups of characters rotated so that each group appeared in all three study conditions. An additional 18 characters were used as fillers appearing at the beginning and end of each study list.

**Procedure.** Participants were tested individually. At the beginning of the study session, the participants were told that the purpose was to test their knowledge of Chinese. They were not informed about memory tests. Each participant studied three separate lists of characters, which were presented on a computer screen. Each list contained 10 target characters, preceded and followed by 3 filler characters. Items were presented one at a time in a random order.

The participants were required to complete all three study conditions. In the physical condition, they were asked to count the number of horizontal strokes in the presented character. In the read condition, they had to read the presented character aloud. In the generate condition, participants were instructed to generate a meaningful phrase from the presented character and were asked to say the answer aloud. All the responses were tape-recorded. The order of the three study conditions was counterbalanced across participants.

On each trial, one of the characters appeared at the center of the computer screen. After making a response, the participant proceeded to the next trial at his or her own pace. After the study session, a distractor test was given. The participants completed a questionnaire written in English, which took about 10 min to complete.

After finishing the questionnaire, 24 participants were given the (implicit) character stem completion test. The 30 test cues were printed on a separate sheet of paper. All test cues were listed on one single page. They could be completed with the target characters that the participants had studied under one of the three study conditions, or with other characters. The participants were instructed to complete each character with the first Chinese character that came to mind. The stem completion test was self-paced, but the participants were asked to finish the test as quickly as possible.

For the explicit memory test, another group of 24 participants was asked to use the stem as a cue to recall the characters that they had studied and to write them down on a blank sheet. Finally, the 24 baseline participants were asked only to complete the stem with the first Chinese character that came to mind. They did not do any of the study conditions.

**Results.** The percent correct for the stem completion test and stem cue recall test as a function of study condition and type of test are presented in Table 2. A 3 (physical vs. read vs. generate study condition) × 2 (stem completion vs. stem cue recall test) analysis of variance was conducted. The main effect of test type was significant \([F(2,29) = 28.08, MS_e = 185.57, p < .001]\). The interaction was also significant \([F(2,29) = 11.60, MS_e = 185.57, p < .001]\).

Tukey HSD tests revealed that the percent correct for the stem completion test was significantly better in all three study conditions, but the percent correct for the stem cue recall test was significantly better than the performance of the baseline group (all \(ps < .001\)), indicating that priming occurred in all implicit testing conditions. There were also significant differences between the physical study condition and the read and generate study conditions (all \(ps < .05\)), but there was no difference between the read and the generate study conditions.

For the stem cue recall test, the generate study condition was better than the read study condition (\(p < .001\)), and the read study condition was better than the physical study condition (\(p < .05\)).

**Discussion.** A typical LoP effect was found in the explicit test; whereas no such effect was found in the stem completion test, in which there was no difference between the read and generate study conditions. Thus, these results do not support the conceptual processing hypothesis or voluntary contamination hypothesis. All the three study conditions differ in the magnitude of the LoP effect. The voluntary contamination hypothesis cannot explain these results, because the present study met the retrieval intentionality criterion, which suggests a dissociation between explicit and implicit tests when all conditions except instruction are held constant. The present finding of differences in the stem completion between the physical and the other two study conditions is best explained by a lack of lexical processing in the physical condition.

However, priming was found in the physical study condition. This was not predicted by the lexical processing hypothesis. One possible reason was that some participants might have processed the target characters as lexical units, whereas others might not have. Since there was no direct and independent index of lexical processing in the physical study condition, there was no way to prevent participants from processing the target character as a lexical unit, even though it was not required by the task. As a result, the overall performance in the physical study condition was worse than both the read and generate study conditions but still showed a significant priming effect.
EXPERIMENT 2

In Experiment 2, we again examined the effect of retrieval intentionality (i.e., test instruction) on the character stem completion test, this time using a different set of LoP manipulations. This was done to test the generality of our conclusion that no LoP effects were present in the character completion test. In this experiment we also introduced the factor of cue type: The stem cue used for testing was either semantically or graphemically related to the character studied.

Method

Participants and Design. One hundred and twenty-five Chinese University of Hong Kong undergraduates participated in Experiment 2. They were all native Chinese (Cantonese) speakers who also spoke English, and all were paid for participating.

The experiment was a mixed factorial design with two between-subjects factors: test instruction (stem completion vs. stem cue recall test) and test cue (semantic vs. graphemic cue) and one within-subjects factor: the study condition (physical vs. phonetic vs. semantic). One hundred participants were randomly assigned to one of the 4 groups. The other 25 participants were assigned to the baseline condition. The participants in the baseline condition performed only the stem-completion test without studying the words.

Materials. The materials were 60 horizontally structured Chinese characters, 30 of which had been used in Experiment 1. All 60 characters consisted of left and right radicals. None shared the same left (semantic) radical. The testing materials for the stem completion test were characters with the right radical removed. There were several possible ways to complete each stem.

In a pilot study, participants who did not take part in the present experiment were asked to judge the relatedness of meaning between each target character and its left radical (stem). Thirty of these stems were judged to have meanings that were different from the meaning of the overall character. These stems were considered to be graphemic cues, because the stems overlapped with the target character only at the graphemic level. The other 30 target characters, also used in Experiment 1, were judged to have meanings that were very similar to the meanings of their stems. These stems were considered semantic cues.

Whether a stem was considered a semantic cue or a graphemic cue depended solely on the target character (see the examples in Table 1). Thus, the participants in both the semantic cue and graphemic cue conditions, for both the stem completion and the stem cue recall test, received exactly the same radicals as test cues. In other words, all participants, including those in the baseline conditions, received the same test cues.

As in Experiment 1, each set of 30 characters was further divided into three groups, which were used in the three study conditions. This time, though, the study conditions were physical, phonetic, and semantic. The three groups of characters rotated so that each group appeared in all three study conditions. Six additional characters were used as fillers, appearing at the beginning and end of each list.

Procedure. The participants were tested in groups of four. They received one of two sets of target characters, depending on the test cue condition to which they were assigned (semantic or graphemic). At the beginning of the study session, the participants were told that the purpose of the study was to test their knowledge of Chinese. They were not informed about memory tests. The participants studied three separate lists of characters, which were presented by means of an overhead projector. Each list of characters contained 10 target characters, which were preceded and followed by three filler characters. Items were presented one at a time in a random order.

The participants were required to complete all three study conditions. In the physical condition, the participants were asked to count how many enclosed areas (“holes”) were in the character.

In the phonetic condition, the participants had to find the first letter of the pronunciation of the Chinese character given in Pinyin, the Chinese alphabetic system. For example, the correct letter for the character “破” (pronounced as “po” in Pinyin) is “p”.

In the semantic condition, the participants were instructed to judge whether the character could be used as a noun (a Chinese character typically has multiple meanings, and its meaning and case are mainly determined by the context). The participants wrote their answers on a piece of paper. The order of the study conditions and the characters used for three study conditions were counterbalanced across participants.

On each trial, one of the target characters was shown to the participants for 2 sec. The participants were allowed an additional 3 sec to make a response. After the study session, the participants took approximately 10 min to complete a questionnaire written in English (this served as a distractor task). After they had finished the questionnaire, the participants were given the character stem completion test or the stem cue recall test.

For the character stem completion test, the 30 stem cues were printed on a single sheet of paper. They could be completed with the target characters that the participants had studied under one of the three study conditions, or with a number of other characters. The participants were instructed to complete each character with the first Chinese character that came to mind. For the stem cue recall test, the participants were asked to recall the characters that they had studied, using the stem as a cue, and to write the characters down on a blank sheet.

The baseline participants did not participate in any of the study conditions. They were asked only to complete the stem with the first Chinese character that came to mind. Each participant in the baseline condition was scored on two different sets of target characters, one for the semantic test cue and one for the graphemic test cue.

<table>
<thead>
<tr>
<th>Type of Cue</th>
<th>Study Condition, and Type of Test in Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic Cue</td>
<td>Physical</td>
</tr>
<tr>
<td>Stem completion</td>
<td>&quot;*Corr.&quot;</td>
</tr>
<tr>
<td>Baseline</td>
<td>15.93</td>
</tr>
<tr>
<td>Priming</td>
<td>10.00*</td>
</tr>
<tr>
<td>Stem cue recall</td>
<td>24.66</td>
</tr>
<tr>
<td>Note: Priming = stem completion. Baseline.</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Percentage Correct as a Function of Type of Cue, Study Condition, and Type of Test in Experiment 2

Type of Cue

Semantic Cue

Physical | "*Corr." | 29.23 | 39.73 | 42.04 | 8.70 |

Graphemic Cue

Physical | "*Corr." | 29.23 | 39.73 | 42.04 | 8.70 |

Semantic | 15.93 | 18.93 | 1.26 |

Priming | 10.00* | 21.18 | 20.89* | 24.41 | 22.11* | 4.09 | 3.08* | 1.85 | 6.60* | 3.63 | 16.30 |

Stem cue recall | 24.66 | 3.66 | 20.89 | 3.40 | 59.30 | 3.31 | 25.40 | 3.36 | 19.40 | 3.63 | 15.00 |

Note: Priming = stem completion. Baseline.
Results

The mean percentages of correct responses for the stem completion test and stem cue recall test as a function of study conditions and types of test cues are presented in Table 3. Priming effects were calculated by subtracting baseline from correct percentages of the stem completion test. For the character stem completion (i.e., the implicit) test, planned comparisons (t tests) revealed that the three study conditions for both test cues were significantly different from the baseline (all ps < .05) and thus yielded significant priming effects.

A 2 (stem completion vs. stem cue recall instruction) × 2 (semantic vs. graphic cue type) × 3 (physical vs. phonetic vs. semantic study condition) analysis of variance was conducted on the priming effects. All the main effects were significant [F(1,96) = 84.90, MS_e = 2.95, p < .001; F(1,96) = 67.65, MS_e = 2.95, p < .001; and F(2,192) = 21.86, MS_e = 1.63, p < .001, respectively]. Moreover, all three of the two-way interactions were significant: cue type × study condition [F(2,192) = 48.51, MS_e = 1.63, p < .001], cue type × testing condition [F(1,96) = 6.91, MS_e = 2.95, p < .001]; and testing instruction × study condition [F(2,192) = 4.99, MS_e = 1.63, p < .001]. Finally, the three-way interaction was also significant [F(2,192) = 15.73, MS_e = 1.63, p < .001].

Tukey HSD tests revealed that for the stem completion test there were no differences in priming between the three study conditions when the cue type was graphic. When the cue was semantic, however, there was a significant difference between the physical condition and the other two study conditions (p < .05 and p < .01, respectively).

On the stem cue recall test, when the cue type was semantic, the participants in the semantic study condition performed better than those in the phonetic (p < .05) and physical study conditions (p < .01). Performance in the phonetic study condition was better than that in the physical study condition (p < .01). When participants received the graphic cue, physical study performance was better than that for semantic study (p < .05). There was no difference between the phonetic and semantic study conditions, or between the physical and phonetic study conditions.

Discussion

The most important finding in Experiment 2 was that there was no difference between the phonetic and semantic conditions in the stem completion test, for either the semantic cue or graphic cue conditions. In other words, no LoP effect was found. However, the semantic cue and graphic cue testing conditions yielded different patterns of results. For the semantic cue condition, priming was smaller in the physical condition than in the phonetic and semantic conditions. This is consistent with the results of Experiment 1, and it provides further support for the lexical processing hypothesis.

On the other hand, when participants were presented with graphic cues, there was no difference in priming among the three study conditions. As mentioned before, a character stem completion test with graphic cues is a perceptual implicit test, whereas a completion test with semantic cues is closer to a conceptual implicit test (the present test differed from other conceptual implicit tests in that semantic cues also overlapped perceptually with the target character). Our result from the graphic cue condition is consistent with the general failure to find LoP effects in perceptual implicit tests in the literature. The semantic cue condition showed much larger priming in both the phonetic and semantic study conditions. This was expected, because the test cue provided not only perceptual information, but also semantic information. As we discussed above, the smaller priming in the physical study condition may have been due to the fact that some participants might not have processed the target characters as lexical units.

The results from the stem cue recall test are consistent with the transfer-appropriate processing account (Baxton, 1989; Roediger, 1990; Roediger & McDermott, 1993; Roediger et al., 1989). Performance was best when there was no overlap between the study and testing conditions. Semantic cues were most useful for the semantic study condition, whereas graphic cues were most effective in the physical study condition. This theoretical account will be discussed in more detail in the General Discussion section.

In conclusion, the present results suggest that there were no LoP effects in either type of character stem completion test. The smaller priming observed in the physical study condition with semantic test cues was probably due to a lack of lexical processing in that condition.

GENERAL DISCUSSION

One important result found consistently across these experiments was that priming in the character completion test did not differ between the semantic and phonetic study conditions. In parallel with Richardson-Klavehn & Gardiner's (1998) explanation, this suggests that there is no LoP effect in the character completion test. The differences between the shallow study (the physical study condition) and other conditions occurred because the shallow condition did not promote lexical processing. Thus, these results are most consistent with the lexical processing hypothesis. We conclude that LoP effects found in the character completion task were not due to the effect of level of processing per se (i.e., a change in the amount of conceptual processing affects memory performance). Instead, the LoP effect was due to a lack of lexical access in the physical condition. Although similar results and arguments have been presented in other studies (Challis & Brodbeck, 1992; Richardson-Klavehn & Gardiner, 1998; Roediger & McDermott, 1993; Weldon, 1991), the present study extended the findings to the Chinese character completion task with varying types of test cues.

Unlike in the character completion test, under the same study and test conditions LoP effects were found in the cue recall test, consistent with previous research (Craik,
Moscovitch, & McDowd, 1994; Graf & Mandler, 1984; Jacoby & Dallas, 1981). This suggests that retrieval tasks involving deliberate recollection are sensitive to the LoP manipulation. In addition, our results across different study and testing conditions are best accounted for by the idea of transfer-appropriate processing (Blaxton, 1989; Roediger, 1990; Roediger & McDermott, 1993; Roediger et al., 1989). This differs from implicit memory tests: When intentional learning and retrieval occur, participants can adopt an optimal strategy for the task (also see Challis, Velichkovsky, & Craik, 1996). As a result, memory performance may be determined by both study and test conditions, as well as the compatibility between the information encoded and the retrieval processing.

The use of Chinese characters allowed us to generate a conceptual implicit test and a perceptual implicit test with the same test cues (i.e., the semantic vs. graphemic cue manipulation). The results from both kinds of implicit tests were consistent with those of previous studies done with conceptual and perceptual implicit tests and supported the lexical-processing hypothesis.

The second important implication of the semantic cue versus graphemic cue manipulation is that it provided a chance to examine the effect of retrieval/test cues on priming, as well as the LoP effect in implicit tests. Studies have suggested that the compatibility of processing operations used at study and at test are important (see Carroll, 1989). For normal adults, Jacoby (1983a) found that re-instating list context affected perceptual identification following the learning of single word lists. In addition, generation effects have been found in word identification (Toth & Hunt, 1990) and in word fragment completion (Gardiner, 1988) only when test items have been presented in fragmented form identical to the study cue. These studies suggest that reinstatement of specific perceptual features (operational specificity) is important for the recruitment of conceptual processes in an implicit perceptual test.

The second line of evidence for this compatibility notion comes from amnesic patients. It has been suggested that the dissociation between explicit and implicit memory results from differences in degree of environmental support and operations required. Implicit memory can become insensitive to a prior experience when the required operations at study and test do not match. In other words, remembering or priming is best explained by the overlap in the types of processes engaged at study and at test (e.g., Jacoby, 1983a, 1983b; Kolers, 1985; Kolers & Roediger, 1984). For example, Schaeter (1985a, 1985b) suggested that priming was critically dependent on the provision at test of a part of the stimulus that has a unitized memorial representation. Richardson-Klavehn and Bjork (1988) found deficits among amnesics when the unit of prior processing was different from that encountered at test. Our results, even though we did not find an LoP effect, further suggest that the semantic information carried by a test cue provides a priming effect in addition to the effect produced by the perceptual information in the same test cue.

Using the word stem completion test with incidental, intentional, or inclusion instructions and measuring both response times and accuracy, Richardson-Klavehn and Gardiner (1998) showed that an LoP effect on an implicit perceptual test was not due to contamination by voluntary retrieval or to sensitivity to conceptual processing. Instead, their results support the lexical processing hypothesis, suggesting that a lexical processing deficit during shallow study processing is critical for interpreting the LoP effect. They used a perceptual implicit test. We not only replicated their main conclusion, but obtained the same findings in a compatible conceptual implicit test (the semantic cue test condition).

In the physical study condition, typically a subletter or subcharacter feature task, the lexical information necessary for completion of the implicit retrieval task may not be encoded. In other words, the participants may not process the characters as lexical units. This may explain why in some cases (e.g., two unreported experiments from our lab), the physical study condition does not even produce a priming effect. However, the physical task does not completely prevent characters from being perceived as lexical units, but merely attenuates lexical processing and reduces the likelihood of such perception's occurring. Thus, in some conditions (e.g., the present two experiments), the priming effect occurs. In order to establish the relationship between the existence of lexical processing and the LoP effect directly, it would be useful to develop an independent measure of lexical processing in future research.

Using compound words as targets (e.g., *jubilant*), Reinitz and Demb (1994) found that priming in perceptual identification occurred only for studied words, whereas priming in word fragment completion occurred for studied and recomposed words. In English compound words, both parts of the compound are themselves words. In our study, the radicals, which combined into a character, were not themselves words even though they carried either the semantic or the phonetic information of the character. Two-character words, which correspond to English compound words, seem more appropriate for examining the Chinese version of the word completion task. In this task, participants would be asked to complete a target word with a character.

In spite of the differences between English word stem completion tests and Chinese character completion tests, the principles of LoP effects and transfer-appropriate processing can be applied to both. In the present study, we not only replicated the basic finding, but also generalized the principles, applying them to a set of compatible perceptual and conceptual implicit tests, that differed from traditional word completion tests in a significant way.

REFERENCES