The retrieval of controlled and automatic aspects of meaning on direct and indirect tests

DOUGLAS L. NELSON, THOMAS A. SCHREIBER, and PATRICIA E. HOLLEY
University of South Florida, Tampa, Florida

The literature concerning implicit memory presents conflicting evidence on the importance of meaning in recovering recently studied words. When the *same cues* are used during testing, indirect instructions reduce levels of processing effects relative to those obtained with direct instructions, suggesting that meaning is not as likely to be retrieved on indirect tests. However, with certain cues, meaning set size of the studied words affects performance even under indirect instructions, suggesting that meaning is retrieved on such tests. The purpose of the present experiments was to resolve this apparent inconsistency. In Experiment 1, the effects of levels of processing and meaning set size were evaluated under direct and indirect test instructions, with the use of stem and word-fragment cues. In other experiments, beginning and ending stem cues were compared, and levels, set size, and instructional effects were evaluated using meaning cues. The findings indicated that levels effects were determined more by test instructions than by test cues, and that set size effects were determined more by test cues than by test instructions. Implications are discussed for transfer-appropriate processing viewpoints and for a model in which it is assumed that performance is determined by searching either explicit or implicit memories.

Researchers have become interested in contrasts involving direct and indirect tests of memory (e.g., Graf & Mandler, 1984; Graf & Schacter, 1985; Richardson-Klavehn & Bjork, 1988; Roediger, 1990; Schacter, 1987). In direct tests, subjects are instructed to recall or recognize recent experiences; in indirect tests, subjects are asked to perform some task without being told to remember specific events. Subjects given the two types of tests encode the material under the same conditions, but the testing procedure for indirect tests makes no reference to this experience and often includes manipulations designed to keep the subjects from consciously thinking about it. In one type of indirect test, for example, subjects are given word stems (e.g., ork) or word fragments (e.g., _ORK) and are asked to produce the first word that comes to mind that fits each cue. To disguise the indirect nature of the task, many test cues are unrelated to words presented as part of the study list.

Interest in the contrast between direct and indirect tests has arisen because many variables have dissociated effects on performance in the two types of tests (e.g., Richardson-Klavehn & Bjork, 1988; Roediger, 1990; Schacter, 1987; Tulving, Schacter, & Stark, 1982). In indirect tests, for example, older adults and amnesics fail to show the deficits they show in direct tests (Light & Singh, 1987; Squire, Shimamura, & Graf, 1985). In addition, for college students, although phonemic interference effects are obtained with word-fragment cues, meaning-related interference effects are not (Nelson, Keelean, & Negrao, 1989). Similarly, although large levels of processing effects are obtained in direct tests such as recall or recognition, these effects are sometimes found (e.g., Hamann, 1990; Srinivas & Roediger, 1990) and sometimes not found in indirect tests (Graf & Mandler, 1984; Graf & Ryan, 1990; Roediger, Srinivas, & Weldon, 1989; Squire et al., 1985).

The reduction of levels of processing effects in indirect tests serves as a primary focus for the present experiments. This reduction has been observed for both stem and fragment cues, and this finding led Roediger and his colleagues to suggest that similar processes are involved for both cues in such tests (Roediger, 1990; Roediger et al., 1989). They classify stem and fragment cues as primarily *data-driven*, meaning that, under indirect conditions, these cues are used to recapitulate data-driven or nonsemantic encoding operations executed during study. This view assumes that remembering is facilitated to the extent that the mental operations performed during study overlap with those performed at test. This view is known as transfer-appropriate processing (TAP), and it explains why levels of processing effects are small or are not obtained with data-driven cues. Meaningfully related information is presumably encoded during study but is not recovered during test because stems and fragments are more transfer-appropriate cues for data-driven operations.

However, the findings also show that such cues become more transfer-appropriate for conceptually driven operations under direct as opposed to indirect test instructions. Roediger and his colleagues found levels of processing effects for both stem and fragment cues when direct test instructions were used (Roediger et al., 1989). Hence, levels of processing effects for fragment and stem cues depend on the directness of the test instructions. The retrieval processes involved in the use of data-driven cues somehow change when subjects are given direct instead of indirect test instructions.

Other findings indicate that meaning plays an important role in stem completion tests, regardless of the directness of the test instructions (Nelson, Canas, Bajo, & Keelean, 1987; Nelson et al., 1989). Evidence for this conclusion comes from manipulations of the meaning set size of the studied word or target, with meaning set size defined as the number of associates linked to the target; for example, the target GEM produces *diamond, jewel, stone,* and so forth. The presence of set size effects in any task has been attributed to the results of search processes involving a target's associates (Nelson, Schreiber, & McEvoy, 1992). Given word stems as test cues, target recovery is more likely when the *target* has a small set of meaningfully related associates than when it has a relatively large set. In contrast, however, set size effects have not been found for word fragments. Meaning set size effects are found for stem cues but not for fragment cues, and this dissociation is apparent under both direct and indirect instructions (Nelson, Canas, et al., 1987; Nelson et al., 1989).

The set size findings stand in apparent contrast with levels of processing findings. Levels effects interact with the directness of test instructions but not with the type of cue, and meaning set size effects interact with the type of cue but not with the test instructions. These apparent inconsistencies and other findings (e.g., Weldon, 1991) make it difficult to determine just how data-driven fragment and stem cues really are. However, the inconsistencies may be more apparent than real if it is assumed that levels of processing effects and meaning set size effects represent two different aspects of meaning. Levels of processing effects may reveal the influence of controlled processing activities performed by the subject during the study trial (Craik & Tulving, 1975). One subject may rate a stimulus for concreteness, whereas another may be asked to name or count its vowels, and the superiority of the semantic activity may appear whenever a direct retention test is used, regardless of whether the test cue is a stem, a fragment, or a meaningfully related associate.

In contrast, meaning set size effects linked to the studied target word may reveal the influence of automatic activation processes. The presentation of a familiar word theoretically activates its closest associates, regardless of whether subjects are rating concreteness or naming vowels, and this activation may be used to support target recovery under a variety of conditions. Previous findings have shown that presentation rate, levels of processing, and the relatedness of interfering lists have large effects on

the probability of target recovery without influencing the magnitude of the target set size effect (Nelson, McEvoy, Janczura, & Xu, 1992; Nelson, Schreiber, & McEvoy, 1992). Target set size effects are apparent even when the targets have been encoded in the presence of rhyming words, suggesting that the closest associates of a familiar word are automatically activated (Nelson, Bajo, & Canas, 1987). However, although set size effects are uninfluenced by a variety of encoding conditions, they are influenced by the nature of the test cue. Set size effects are obtained with stem, rhyme, and associatively related test cues, but they are not obtained with fragment cues; nor are they obtained when the target itself is used as the cue in a typical recognition test (Nelson, Canas, & Bajo, 1987).

These considerations suggest that directness of the test instructions affects the probability of retrieving meaning encoded as a result of controlled processing activities (e.g., levels of processing), whereas the nature of the test cue affects the probability of retrieving meaning activated as a result of more automatic processing (e.g., set size). The controlled–automatic processing distinction has been useful for analyzing a number of issues (e.g., Hasher & Zacks, 1979; Shiffrin & Schneider, 1977), and this distinction represents a crucial aspect of a model for describing the relationship between explicit and implicit memories, called PIER (Nelson, Schreiber, & McEvoy, 1992). Although too complex to be presented in detail, the most important assumptions for our present purposes can be summarized briefly.

PIER assumes that two types of representations are produced as a result of the study experience. One type is related to the nature of the controlled processing activities—that is, to the nature of what subjects do or are instructed to do during study—including rehearsal of list words, levels of processing, imagery, and so forth. Intentional processing activities produce an explicit memory representation. The other type of representation is related to the activation of memories connected to the word prior to the laboratory experience. PIER assumes that each word's meaningfully related associates are automatically activated and that this activation produces an implicit memory representation. Explicit and implicit representations of the target are produced, respectively, as a result of controlled and automatic processes during the encoding phase of the task. Hence, in PIER, explicit and implicit memories refer to representations produced as a result of different encoding activities, and the directness of the test instructions refers to the subjects' orientation toward the study trial experience during testing. Following Richardson-Klavehn and Bjork (1988), PIER uses different terms to refer to hypothetical memories and methods of measuring memory performance in order to highlight important differences between these concepts.

PIER also assumes that targets can be recovered either as a result of searching through explicitly created memories or as a result of searching through implicitly activated memories. The two search processes are related by exclusion (see Jones, 1987). In cuing tasks, the model as-

sumes that retrieval is initiated by the presentation of the test cue with the use of the cue determined by the test instructions. Direct instructions refocus attention on controlled processing activities carried out during study and theoretically increase the probability of searching explicit memories. In contrast, indirect instructions decrease the probability of searching explicit memories. Although directness of the test instructions affects the likelihood of searching explicit memories, directness has no effect on the likelihood of searching implicitly activated memories. Furthermore, the explicit search should be more likely with direct instructions, regardless of whether the cues consist of stems, fragments, rhymes, or meaningfully related associates. As a result, PIER predicts that levels of processing effects should be more apparent for direct than for indirect instructions and that this interaction should be apparent with a variety of test cues.

Target set size effects, in contrast, arise as a result of sampling memories implicitly activated by the test cue and by the target. The test cue activates representations in the domain of information specified by the cue and the instructions governing its use. For example, as evidenced by comparable lexical set size effects, both fragment and stem cues activate lexical representations of related words that share letters and sounds with the test cue (Nelson, Canas, et al., 1987). In PIER, the sampling of a lexical representation related to the test cue can be used in one of two ways, depending on the characteristics of the cue (Nelson, Bajo, & Canas, 1987; Nelson, Canas, et al., 1987; Nelson et al., 1989). This representation can be used to produce the name of the target directly, or, alternatively, the lexical representation can be used as an internally generated cue to initiate search processes involving meaningfully related associates of the target. Targets can be recovered either as a result of sampling lexical representations activated by the test cue or as a result of using these representations to initiate searches through associates of the target implicitly activated during study. For *fragment* cues, lexical representations are recovered directly as names. Regardless of whether test instructions are direct or indirect, the instructions require the production of a word that fits the constraints imposed by the letters and by the spaces provided for missing letters. According to PIER, the constraints produced by the spaces for missing letters limits the search to low-level lexical representations and curtails the search through meaningfully related associates. As a result, meaning set size effects linked to the target should not be apparent for fragment cues.

For *stem* cues, spaces for missing letters are not presented as part of the cue and sampled lexical representations are more likely to be used to initiate searches within the meaning domain. The meaning search should be especially likely when a stem cue is only weakly related to its target, making it difficult to produce the target directly. As a result, meaning set size effects linked to the target should be apparent for weak stem cues. In this account, the important difference between fragment and stem cues

does not arise from differences in meaning inherent in the cues themselves but from the presence of spaces for missing letters. The presence of spaces changes the focus of the search from one involving meaning to one based on recovering a word name that matches the letter requirements of the cue itself. Hence, PIER predicts that meaning set size effects should interact with the type of test cue, but not with levels of processing manipulations or with the directness of the test instructions. Meaning set size effects should not interact with levels of processing manipulations because the two effects are mediated by mutually exclusive search processes, and, for the same reason, set size effects should not interact with the directness of instructions. Such instructions only affect the likelihood of searching explicit memories.

The purpose of Experiment 1 was to evaluate several of these expectations. Directness of the test instructions was crossed with levels of processing during an encoding phase and with type of test cue (either weakly related word stems or fragments). In addition, meaning set size of the target words was varied so that half of the words had small and the other half had large sets of related associates. Three predictions were made by PIER: (1) Levels of processing effects should be apparent for both stems and fragments under direct but not under indirect instructions. Such a result would be consistent with the assumption that test instructions affect the likelihood of searching memories created as a result of explicit processing activities. (2) Meaning set size effects should be apparent for weak stem cues but not for weak fragment cues. This result would indicate that the more automatic aspects of meaning are controlled by type of test cue. And (3) levels of processing and meaning set size, as well as directness of the test instructions and meaning set size, should have additive effects on target recovery. These findings would be consistent with the assumption that the search of explicit memory and the search of implicit memory are mutually exclusive.

Each of these predictions has been investigated and confirmed in separate investigations (Nelson, Bajo, & Canas, 1987; Nelson, Canas, et al., 1987; Nelson et al., 1989; Nelson, Schreiber, & McEvoy, 1992; Roediger et al., 1989). The purpose of Experiment 1 was to bring this set of divergent predictions together within the confines of a single experiment because the predictions might fail as a result of the control exerted over variables that were left uncontrolled in the separate investigations. For example, in our studies with stem and fragment cues, levels of processing were not manipulated and could have varied substantially between subjects whenever we used intentional learning instructions. In Roediger's studies (Roediger et al., 1989), meaning set size was not manipulated and, although unlikely, it is conceivable that levels effects were observed only for words having limited meaning sets. When both types of semantic processing are carefully controlled, the expected effects might not emerge. Finally, an important characteristic of the research designs used in the present experiments is that manipula-

tions of the directness of the test instructions were always crossed with manipulations of type of test cue. In earlier studies, these variables were often confounded, making it difficult to draw unambiguous conclusions about the relationship between these variables (e.g., Blaxton, 1989; Srinivas & Roediger, 1990; Tulving et al., 1982).

The purpose of Experiments 2 and 3 was to vary stem-to-target strength to test the prediction that strong stems, in comparison with more weakly related stems, should reduce the magnitude of the meaning set size effect. Finally, the goal of Experiment 4 was to test PIER's prediction that meaningfully related test cues will produce the same pattern of effects as will stems in terms of the predicted effects of levels of processing, meaning set size, and directness of the test instructions.

EXPERIMENT 1

Method

Design. The design formed a 2^4 mixed-model factorial. Directness of the test instructions (direct, indirect), levels of processing during study (name vowels, rate concreteness), and type of test cue (fragment, stem) were manipulated between subjects. Target set size (small, large) was manipulated within subjects.

Subjects. Sixteen subjects were randomly assigned to each between subject condition, so that a total of 128 subjects participated in the experiment. All subjects were selected from courses in introductory psychology and received course credit for participation.

Materials. Normative estimates of target set size were available for 2,500 words, and these estimates were used in constructing two lists for each condition as shown in Appendix A. Set size was measured by presenting words to groups of students (n = 120–180), who responded by writing the first word that came to mind for each presented word (see, e.g., Nelson & Schreiber, 1992, for details). The number of different words provided by 2 or more subjects in the normative sample was used to estimate set size. As can be seen, 24 words served as targets in each list. Each had been a stimulus in the normative task, with half defining small and half defining large sets of associates (e.g., GEM had a small set of 7 associates and FADE had a large set of 23 associates in the norms). Targets with small sets averaged 6.71 associates (SD = 1.27), and those with large sets averaged 19.75 associates (SD = 2.38). Word concreteness was equated within each condition of set size and averaged 5.12 (SD = 1.12).

Word stems (e.g., EM, OG, USK, etc.) and word fragments (e.g., _EM, _OG, _USK, etc.) were determined for each target by independent norming procedures. Available norms were used for word stems that consisted of the last few letters of the target, beginning with the vowel. After hearing the stem, subjects wrote the first word to come to mind that rhymed with it (Nelson & McEvoy, 1979). This auditory procedure does not appreciably affect the results relative to a visual presentation (Nelson, 1981). Word-fragment norms for the targets were obtained by deleting letters from the targets and replacing them with spaces. The cues were presented in booklets, and the subjects were asked to produce the first word to come to mind that fit the cue. The cue set sizes of stems and fragments were determined by counting the number of different but appropriate words. These values were equated for each type of cue within each level of target meaning set size, and across all conditions set size averaged 11.02 (SD = 6.56). Relative cue-to-target strength was determined by calculating the probability of each target in relation to its cue and, as with cue set size, strength was equated in all conditions and averaged 0.14 (SD = 0.11). Hence,

each cue would be expected to produce about 11 words, and the cues would be expected to produce their targets about 14% of the time in the absence of a study phase. The majority of the fragment cues (77%) were identical to the ending stem cues and consisted of the same stems with spaces indicated for missing letters.

Procedure. Subjects participated in single sessions and received a single study–test trial. Target words appeared in uppercase letters and were presented by a Kodak carousel projector at a 3-sec rate. Subjects given concreteness rating instructions were told that the experimenter was interested in word meaning, particularly in differences between words in the degree to which they refer to objects. Subjects were asked to rate each word for concreteness on a 5-point scale with 1 = very abstract, 5 = very concrete. Examples were provided to help subjects understand the task. Subjects given the vowel instructions were asked to locate the vowels in each word and to report them aloud as rapidly as possible. Both instructional groups were shown a series of slides containing a row of Xs to acquaint them with the presentation rate just before the word list was shown.

As soon as the last word was presented, the experimenter read the testing instructions. Subjects given direct instructions were told that their memory for the words just studied would be tested and that a word stem (or word fragment) would be presented for each studied word to aid their attempt at recall. They were also told that only half of the cues were related to studied words, and that if a word from the study list could not be recalled, they were to generate the first word to come to mind that fit with the letters available in the cue. Subjects given indirect instructions were first asked to complete stems (or fragments) for several city names to acquaint them with the procedure. After completing this task, they were asked to complete each stem (or fragment) with the first real word that came to mind that fit the cue. For these subjects, memory was never mentioned in any phase of the experiment. For all subjects, each cue appeared in uppercase letters, with spaces indicated for missing letters in the fragment cues, and each cue was individually presented at a subject-paced rate. The cues for Lists 1 and 2 were presented to all subjects and were randomly intermixed. Hence, subjects were shown words from either List 1 or List 2 during the study phase, but they were tested on the cues from both lists. This procedure allowed us to establish a baseline completion rate for cues from the unstudied list within the confines of the experiment. This rate can differ somewhat from the rate determined by the norms, because fewer subjects are involved in the experiment than in the normative sample.

Results and Discussion

The probability of recovering targets is shown in Table 1 for each condition. The bottom half of this table shows the baseline measures, and the top half shows the measures obtained for the words presented during the study trial. A separate analysis of variance performed on the baseline measures indicated that none of the four factors represented in the table were significant, nor were any of the interactions. The criterion for significance was set at .05 in this and in all other analyses. The probabilities of completing stem and fragments cues for unstudied words were, respectively, .22 and .20. Although the completion rates were somewhat higher than what was expected on the basis of the normative assessment, this analysis indicated that the completion rate did not differ appreciably across conditions, and, as a consequence, base rates were ignored in the analysis of the recovery rates in the analysis of studied words (including unstudied

Table 1
Probability of Target Recovery as a Function of Prior Study, Levels of Processing,
Test Instructions, Type of Test Cue, and Meaning Set Size (Experiment 1)

| | Type of Test Cue | | | |
| | Fragment Cue | | Stem Cue | |
	Small Set Size	Large Set Size	Small Set Size	Large Set Size
Studied Words				
Direct Instructions				
Vowel naming	.51	.47	.49	.48
Concreteness	.66	.62	.70	.56
Indirect Instructions				
Vowel naming	.46	.50	.54	.47
Concreteness	.50	.51	.56	.48
Unstudied Words				
Direct Instructions				
Vowel naming	.21	.20	.24	.33
Concreteness	.17	.26	.24	.20
Indirect Instructions				
Vowel naming	.19	.21	.18	.17
Concreteness	.19	.20	.21	.16

words in the analysis did not alter any of the conclusions made about studied words).

Examination of the probabilities of target recovery for studied words indicated that recovery was more likely under direct (.56) than under indirect (.50) instructions, for concreteness ratings (.57) in comparison with vowel naming (.49), and for targets with small sets (.55) in comparison with targets with large sets (.51). Fragment cues (.52) and stem cues (.53) produced about the same level of target recovery. The results of an analysis of variance indicated that effects of test instructions [$F(1,120) = 8.85$, $MS_e = .024$], levels of processing [$F(1,120) = 18.59$], and set size [$F(1,120) = 8.02$, $MS_e = .012$] were significant, and that differences between the two types of test cues were not ($F < 1$).

The statistical analysis also indicated that the two expected interactions were reliable. First, test instructions interacted with levels of processing [$F(1,120) = 9.66$]. The probabilities of target recovery under direct instructions were .49 and .63, respectively, for the vowel and concreteness tasks, whereas, under indirect instructions, these means were .49 and .53. A Fisher's two-tailed least significant difference (LSD) of 0.05 indicated that the effects of levels of processing were reliable under direct but not under indirect instructions. This interaction replicates previous findings obtained with fragment and stem cues (Roediger et al., 1989). Second, type of test cue interacted with set size [$F(1,120) = 5.99$, LSD $= 0.04$]. Given stem cues, the probabilities of recovery for targets with small and large sets were .57 and .50 and, with fragment cues, these probabilities were .52 and .52. Target set size effects were significant for stem cues but not for fragment cues. Once again, this interaction replicates the pattern found in previous experiments (Nelson, Canas, et al., 1987; Nelson et al., 1989). Finally, interactions between test instructions and set size ($F = 1.22$) and between levels of processing and set size ($F = 2.22$) were not significant,

nor were any of the remaining sources in this analysis significant, with most Fs less than or near unity.

The results of this experiment indicate that directness of the test instructions interacts with levels of processing but not with meaning set size, and that type of cue interacts with meaning set size but not with levels of processing. One implication of these findings is that levels effects and set size effects are produced by different types of mental activities. According to PIER, manipulations of levels of processing influence the nature of the explicit encoding and, as predicted, levels effects should be more apparent when the test instructions or other task conditions encourage search of the representation of the list created as a result of explicit processing activities. Direct test instructions increase the probability of such search relative to indirect instructions, because these instructions refer subjects to the study episode. As a consequence, concreteness ratings facilitate recovery relative to vowel naming only under instructions to use the test cues to recover list words.

In contrast, meaning set size effects are obtained regardless of the test instructions. These effects, however, depend on the nature of the test cue. Set size effects are obtained for stem cues, but they are not obtained for word fragment cues, and this finding is consistent with other results showing that performance often differs for the two types of cues (e.g., Squire, Shimamura, & Graf, 1987). According to PIER, the difference between these two types of cues controls the probability of searching through meaningfully related associates of the target during testing. Theoretically, stem cues activate the name of the target as well as competing names having similar visual-phonemic patterns, and, to facilitate selection, a search of the meaningfully related associates linked to each sampled name is likely to occur. Fragment cues provide a similar portion of the orthographic-phonetic features of the target, at least in this experiment, but these cues also pro-

vide spaces for missing letters, and this difference appears to change the nature of the search process (Nelson, Canas, et al., 1987; Nelson et al., 1989). Instead of searching meaningfully related associates linked to each activated name, the subject samples words as orthographic–phonetic patterns in an attempt to produce a word that fits the constraints provided by the letters and spaces in the test cue. The presence of the spaces for the missing letters appears to alter the nature of the search process.

EXPERIMENT 2

One purpose of Experiment 2 was to determine whether strongly related stem cues would reduce the magnitude of the target set size effect relative to weakly related stems. According to PIER, strongly related stem cues should increase the probability that the name of the target can be recovered directly from the lexical entries activated by the test cue. In this model, the test cue activates related concepts in the domain of information specified by the test instructions and the test cue. For example, stem cues activate lexically related entries sharing the same phonetic features. These entries are then sampled on a probabilistic basis, and the recovery of each entry is used directly as the response or it is used as an internally generated cue to initiate search in another domain such as meaning. Evidence for these search processes can be found in cued recall studies showing that, for stem cues, recall is a function of both lexical set size defined by the test cue and meaning set size defined by the target (see Nelson, Schreiber, & McEvoy, 1992, for a review). These searches presumably proceed in a parallel fashion in that the lexical search may continue while the meaning-related search has been initiated for entries sampled earlier. With strongly related stems, the target is likely to be recovered as a result of the lexical search for a name that meets the task requirements, but with weakly related stems, the target is likely to be recovered only as a result of searching through meaning-related associates. For weakly related stems, recall is more likely to be based on the success of a search through the meaning domain, and, as a result, meaning-related target set size effects should be more apparent for weak than for strong stem cues. This result would be consistent with the general assumption that the presence of meaning set size effects in various memory tasks are determined by the nature of the test cue.

To evaluate these expectations, we manipulated the preexperimental strength of stem cues in Experiment 2 and we crossed this manipulation with target set size. Stem cues were strong or weak, and meaning set size of the target was either small or large. Only direct test instructions were used, because meaning set size effects are unaffected by the directness of the instructions and, for the same reason, levels of processing were not manipulated. Subjects were simply told to remember as many words as possible. In addition to manipulating target set size and stem strength, the stems consisted of either the beginning sound of the target or its ending sound. We included type

of stem as a variable for two reasons. First, researchers sometimes use one type of stem or the other, and it would be useful to know whether it makes a difference in recall when other known characteristics are controlled. Second, although we have obtained target set size effects with ending stems in many experiments, we have never determined whether such effects would also be present with beginning stems, and we thought that including them would provide another test of PIER. This model predicts that, other things being equal, beginning and ending stems should produce the same effects. Target set size effects were expected to be more apparent when the stems were weakly related to their targets than when they were strongly related, and this pattern was expected to be the same for both beginning and ending cues.

Method

Design. The design formed a 2^3 mixed-model factorial with type of stem (beginning, ending) manipulated between subjects and with stem-to-target strength (strong, weak) and target set size (small, large) manipulated within subjects. Twenty subjects served in each between subject condition, with 10 assigned to each list.

Materials. Two different lists were constructed by using the normative data used for the first experiment, and these lists are presented in Appendix B. Targets had small or large associative sets averaging, respectively, 6.58 associates ($SD = 1.67$) and 20.35 associates ($SD = 2.83$). Word concreteness was equated and averaged 4.92 ($SD = 1.29$).

The stem cues used to cue the recall of these targets were selected from norms of beginning and ending sounds in which subjects wrote the first word coming to mind that sounded like or that rhymed with the sound just heard (see, e.g., Nelson & McEvoy, 1984). Ending stems consisted of the end of the word, starting at the vowel, and beginning stems consisted of the beginning of the word, ending at the first vowel. These norms were used to estimate the number of words sharing that sound and the relative strength of each word in relation to the stem cue. For both types of stem cues, the probability that strong cues produced the target words was 0.36 ($SD = .11$), and for weak cues, this probability was .05 ($SD = .03$). Stem-to-target strength was equated for both cues at each level of target meaning set size. Stem-defined set size was also equated within each level of target meaning set size and at each level of strength, but it tended to be slightly smaller for beginning stems, 7.79 ($SD = 2.63$), than for ending stems, 12.06 ($SD = 5.80$).

After the experiment was completed, we discovered that ending stems shared a greater proportion of letters with their targets (71%) than did beginning stems (55%). This problem arises because more letters are typically used in English words from the vowel to the end of the word than from the beginning letter to the first vowel. Although there are many exceptions to this rule, the confounding appears to be natural. Fortunately for us, the proportions of overlapping letters for each type of stem cue were equated at each level of target set size and stem-to-target strength. Hence, for a given stem, the evaluation of the target set size and strength interaction was uncontaminated. This confounding raised a problem only for interpretation of the main effect associated with stem type (this problem is considered separately in Experiment 3).

Procedure. The procedures governing the study trial were identical to those used in Experiment 1, except that subjects were told to remember the words. They were told that their memory for the words would be tested, but they were not told how it would be tested. The test trial instructions indicated that beginnings (or endings) of the studied words would be presented as cues to help the subjects recall the list words, that the experimenter would pronounce each cue, and that the subjects should repeat the sounds and use them

Table 2
Probability of Target Recovery as a Function of Type of Stem Cue,
Stem-to-Target Strength, and Meaning Set Size of the Target (Experiment 2)

Type of Stem Cue	Stem-to-Target Strength				
	Strong Stems		Weak Stems		
	Small Set Size	Large Set Size	Small Set Size	Large Set Size	M
Beginning stem	.73	.74	.71	.53	.68
Ending Stem	.94	.87	.78	.65	.81
M	.84	.81	.75	.59	

to recall their related list words. This test was self-paced, guessing was allowed but not required, and, as in Experiment 1, the order of all study words and test cues was independently randomized for each subject. Unlike in Experiment 1, each stem cue was related to a study word.

Results and Discussion

The probability of correct recall for each condition is displayed in Table 2. As can be seen, ending cues were more effective than beginning cues, strong cues were more effective than weak cues, and targets having smaller associative sets were more likely to be recalled than those with large sets. Most importantly, for each type of stem cue, target set size effects were less apparent for strongly related stem cues than for weakly related stem cues. These observations were supported by the results of an analysis of variance which showed significant effects for type of stem cue [$F(1,38) = 13.69$, $MS_e = .055$], stem-to-target strength [$F(1,38) = 34.04$, $MS_e = .030$], and target set size [$F(1,38) = 14.89$, $MS_e = .025$]. The interaction between stem-to-target strength and target set size was also significant [$F(1,38) = 11.06$, $MS_e = .016$, LSD = .06]. This interaction is displayed in the bottom row of the table, and, as suggested, target set size effects were reliable when recall was cued by weakly related stems but not when it was cued by strongly related stems.

The relative recall advantage for targets with smaller associative sets was apparent for weakly related stem cues but not for strongly related cues. This finding is consistent with expectations derived from PIER, in that it indicates that target set size effects are determined by the nature of the test cue. PIER explains this difference by assuming that strongly related stems are more likely to lead to sampling the lexical representation or name for the target. With weak stems, related words having similar visual-phonemic patterns are activated and sampled just as with strongly related stems. However, because competition from other similar-sounding words at this level prevents unambiguous target selection, these words are used as internally generated cues to initiate searches at the meaning level. Words sampled at this level are used as cues to activate their meaning-related associates, and the search through these associates is what produces meaning set size effects.

EXPERIMENT 3

The findings of Experiment 2 showed that the interaction between stem-to-target strength and target set size was apparent for both beginning and ending stems. This result, which is consistent with expectations based on PIER, suggests that the two types of stem cues function as retrieval cues in similar ways. However, ending stems were considerably more effective as cues than were beginning stems, and, because of the confounding of overlapping letters, it is impossible to determine the origin of the difference. This difference could have arisen from an inherent superiority of ending sounds as retrieval cues, which is doubtful (e.g., Horowitz, White, & Atwood, 1968), or it could have arisen because ending stems shared more letters with their targets than did beginning stems. Although this difference was not critical to the theoretical ideas motivating these experiments, it represented an annoying "loose end," and Experiment 3 was designed to compare the effectiveness of the two cues in the absence of the confounding. Without the confounding, differences between the two types of cues were expected to disappear. Type of stem was crossed with meaning set size of the target in an attempt to replicate set size effects with beginning cues.

Method

Design and Subjects. The design formed a 2 × 2 mixed-model factorial, with stem type (beginning, ending) varied between subjects and with target set size (small, large) varied within subjects. Forty subjects served in the experiment, 20 in each stem condition, with 10 assigned to each list.

Materials and Procedure. The two lists constructed for each stem condition are shown in Appendix C. These materials were selected from the normative database used for the previous experiments. Target set size was equated for each stem type and averaged 6.42 associates ($SD = 1.45$) and 18.85 associates ($SD = 2.50$), respectively, for small and large sets. Target concreteness was also equated in all conditions and averaged 4.90 ($SD = 1.27$). Stem-defined set size, stem-to-target strength, and the proportion of overlapping letters were also equated across conditions averaging, respectively, 12.03 ($SD = 4.97$), 0.14 ($SD = .16$), and 61% ($SD = .07$). The stem cues defined moderately sized sets of phonemically related words, were moderately related in strength to their targets, and shared approximately 2 out of 3 letters with them. The study-test trial procedures and instructions were identical to those used in Experiment 2.

Results and Discussion

For beginning stems, the probabilities of correct recall for targets with small and large sets were .72 and .60; for ending stems, these values were .73 and .65. Target set size effects were apparent for both types of stems, and there appeared to be no differences in effectiveness between the stems. An analysis of variance of these data indicated that the target set size effect was significant [$F(1,38) = 20.78$, $MS_e = .010$] and that neither the effect of type of stem nor the interaction between type of stem and target set size was significant (both $Fs < 1.00$).

The results of this experiment replicate the target set size effects for beginning stems and cues with intermediate strengths, and they indicate that the difference between beginning and ending stems obtained in Experiment 2 was probably due to the confounding of overlapping letters in that experiment, not to some inherent superiority of ending stems. Finally, this experiment, along with the first two, indicates that, as long as weak to moderately related stems are used to prompt target recovery, such recovery will be more likely when the target activates a smaller set of meaningful associates than when it activates a larger set. PIER attributes this advantage to a search of meaningfully related information activated during study and reactivated during test during the retrieval process. The mere presence of target set size effects indicates that stem cues cannot be classified only as "data-driven" cues. Stem and fragment cues produce levels of processing effects under direct recall instructions, and, provided that they are only moderately or weakly related to their targets, stem cues produce target set size effects regardless of instructions. These and related findings suggest that stem cues are neither exclusively data-driven nor exclusively conceptually driven. Stems can serve either or both functions, depending on testing instructions and on how strongly the stems are related to the name for the target.

EXPERIMENT 4

The purpose of Experiment 4 was to test PIER's predictions concerning the relationships between levels of processing, directness of the test instructions, and meaning set size by using *meaning-related test cues*. As with weakly related stem cues, PIER predicts that each of these variables should affect the probability of recovering the target. Relative to the appropriate contrast conditions, recall should be more likely after concreteness ratings, when test instructions are direct, and when target set size is small. The effect of levels of processing, however, should depend on whether test instructions are direct or indirect. According to PIER, directness of the test instructions affects the probability of searching the explicit encoding in such a way that this encoding is more likely to be searched under direct than under indirect instructions. Given that the levels manipulation affects the accessibility of the explicit encoding, levels of processing effects should be more apparent when subjects are told to use the test cue to recall a list word than when they are told to use the

cue to produce the first meaningfully related word that comes to mind.

PIER, however, leads to different expectations about the effects of meaning set size. Because PIER assumes that the explicit and implicit searches are mutually exclusive, meaning set size effects should not vary with either the levels manipulation or the manipulation of the directness of the test instructions. The presence of meaning set size effects theoretically reflects the results of a search process applied to implicitly activated associates of the target, with success determined by the number of items in the activated set. Because manipulations of levels and test instructions theoretically affect only the nature of the explicit encoding or the probability of searching this encoding, respectively, neither of these variables should influence the magnitude of the observed set size effects.

Method

Design and Subjects. The design formed a 2^3 mixed-model factorial in which levels of processing (rate concreteness, name vowels) and test instructions (direct, indirect) were manipulated between subjects. Meaning set size of the target (small, large) was manipulated within subjects. Sixteen subjects served in each between subject condition, with 8 assigned to each list.

Materials. The two related lists presented in Nelson, McEvoy, and Schreiber (1990, Table 3) were also used in this experiment. Half of the 24 targets in each list had small meaning sets, averaging 7.13 associates ($SD = 1.09$), and the remaining half had large sets, averaging 19.34 associates ($SD = 2.38$). The test cues consisted of 48 words, half of which were related and half of which were unrelated to the targets. The 24 related test cues were selected so that they normatively produced their targets with an average probability of .23 in each condition of target set size ($SD = 0..10$)—for example, BEACH was used as a cue for the target word SAND. The 24 unrelated cues were different for each list, were taken from the norms, and were unrelated to any of the targets in their respective lists.

Procedure. The general procedures for the study and test trial followed those used in Experiment 1. Each subject received a single list during study, and unrelated cues were presented during testing to help disguise the purpose of the indirect test. The subjects serving in direct test conditions were told about the test only after the study trial, and those serving in indirect test conditions never heard the words *memory* or *test* at any point in the experiment. For all subjects, the related and unrelated test cues were randomly intermixed throughout testing.

Indirect test subjects were told that we needed word association data for a few words to be used in future experiments. Related cues were accompanied by the instruction "generate," which subjects understood as an instruction to produce the first word to come to mind. Subjects were asked to avoid responding with proper nouns. Unrelated cues were accompanied by the instruction "read," which reminded them to read the presented word. Direct test subjects were told that some words would be presented to help them remember the words of the list just studied. Each test cue was accompanied by a word indicating how it should be used: "Recall" indicated that the cue should be used to recall a related word from the list; "read" told them to read the presented word. We did not ask the direct subjects to free associate to the unrelated test cues to prevent them from having to randomly switch between having to follow direct recall instructions on one cue and free association instructions on another. Such a mix of instructions might have encouraged subjects to free associate even when given recall instructions, thereby reducing the effectiveness of the instructional manipulation.

Table 3
Probability of Target Recovery as a Function of
Levels of Processing, Directness of the Test Instructions,
and Meaning Set Size of the Target (Experiment 4)

Meaning Set Size	Direct Test Instructions		Indirect Test Instructions	
	Concreteness	Vowel Naming	Concreteness	Vowel Naming
Small	.77	.55	.46	.39
Large	.60	.40	.30	.23
Mean	.69	.47	.38	.31

Results and Discussion

Table 3 presents the probabilities of recovering the target words, as a function of the principal variables. As can be seen, target recovery appeared to vary as a function of levels, instructions, and set size. Recovery was more likely after concreteness ratings than after vowel naming (.53 and .39), more likely after direct than after indirect instructions (.58 and .34), and more likely when set size was small rather than large (.54 and .38). As shown in the bottom row of the table, the advantage for concreteness ratings was clearly more apparent after direct instructions than after indirect instructions. An analysis of variance indicated that levels [$F(1,60) = 23.16$, $MS_e = .027$], test instructions [$F(1,60) = 64.32$], and meaning set size [$F(1,60) = 51.92$, $MS_e = .016$] were each significant sources of variance. Moreover, the interaction between levels of processing instructions and the directness of the test instructions was also significant [$F(1,60) = 6.23$, LSD = .08]. Levels of processing effects were significant under direct instructions but not under indirect instructions. No other interactions even approached the criterion for significance (all $Fs < 1$).

These findings are consistent with expectations based on PIER. Target recovery varied with levels of processing, with test instructions, and with set size, and neither test instructions nor levels reliably affected the observed magnitude of the target set size effect. Finally, as predicted, the levels effect was more apparent under direct than under indirect instructions. Concreteness ratings produced a higher level of target recovery when subjects were told to use some of the cues to recall list words than when subjects were asked to use the cues to generate the first words that came to mind.

GENERAL DISCUSSION

The present findings make three main contributions toward understanding the relationship between what has been encoded or activated during study, the type of test instructions, and the type of information available in the test cue. First, the magnitude of levels of processing effects is determined more by the directness of the test instructions than by the nature of the information available in the test cue. Concreteness rating during study was invariably associated with higher levels of target recovery than was vowel naming, and this advantage was more apparent under direct than under indirect instructions. This interaction with test instructions was obtained for word fragments and word stems and replicates the results of

a number of different experiments (e.g., Graf & Mandler, 1984; Graf & Ryan, 1990; Roediger et al., 1989; Squire et al., 1985). In addition, this interaction was also apparent for meaning cues. In general, the magnitude of the levels of processing effect appears to be affected more by the nature of the test instructions than by the nature of the test cue.

Second, the presence of the meaning set size effect appears to be determined more by the information available in the test cue than by the directness of the test instructions. Words having smaller associative sets were more likely to be recovered than were those with larger sets, regardless of whether the test instructions were direct or indirect. Given that the nature of the test cue determines the presence or absence of set size effects, instructions to use the test cues to recall list words do not produce lesser or greater set size effects relative to instructions to produce the first word that comes to mind. Target set size effects, however, vary considerably with the nature of the information represented in the test cue. Set size effects are obtained with ending and with beginning stems that, in the normative sense, are relatively weak cues for the targets. Such effects are also obtained with meaning-related test cues (Experiment 4; Nelson, Schreiber, & McEvoy, 1992). In contrast, such effects are not obtained with fragment cues (Experiment 1; Nelson, Canas, et al., 1987; Nelson et al., 1989), nor are they obtained with stem cues that are strongly related to their targets (Experiment 2). In comparison with levels effects, meaning set size effects appear to be uninfluenced by the directness of the test instructions but depend on the nature of the test cue. This contrast in results suggests that, although both levels and set size effects are linked to meaning, they are linked to fundamentally different aspects of semantic processing.

The third contribution of these and related findings is that meaning set size effects do not appear to be affected by the levels of processing manipulated during study. Meaning set size effects were equally apparent for beginning and ending stems and for meaning-related test cues, regardless of whether subjects were rating concreteness or counting vowels during study. Provided that the test cue elicits such effects, the relative difference in the recovery rates for words with small and large sets remained constant, regardless of how well the target was encoded. These findings replicate and extend previous results (Nelson, Schreiber, & McEvoy, 1992). Manipulations of levels of processing and presentation rate affect the probability of recovering the target without affecting the magnitude of the target set size effect. These findings suggest that the

set size effect is unrelated to how well the target is encoded as a result of *explicit* processing activities during the study trial.

These three findings have implications for TAP approaches to explanations of differences between various retention tests (e.g., Graf & Ryan, 1990; Roediger, 1990). According to the TAP framework proposed by Roediger and his colleagues, dissociations or interactions between direct and indirect tests reflect differences in the overlap between the mental operations applied at study and those applied at test (Blaxton, 1989; Roediger, 1990; Roediger & Blaxton, 1987; Roediger et al., 1989; Srinivas & Roediger, 1990). Conceptually driven tests such as free recall are presumed to benefit most from conceptual encoding operations, whereas data-driven tests such as fragment completion presumably benefit most from perceptual encoding operations. An additional assumption of this version of TAP is that there is no necessary correlation between the level of processing beneficial for a particular test and the directness of the test instructions (Srinivas & Roediger, 1990). Levels of processing effects can presumably occur on both direct and indirect tests. This assumption is particularly useful for explaining dissociations between tests when directness has been confounded with type of test cue (e.g., comparisons of free recall and fragment completion) and when two indirect tests are compared (e.g., comparisons of fragment completion and free association in response to category names). In these comparisons, dissociations are attributed to differences in TAP, not to differences in directness of the test instructions. This version of TAP provides an explanation of many of the findings reported in the implicit memory literature without assuming that different memory systems are involved in different tests (see, e.g., Hayman & Tulving, 1989).

This version of the TAP approach, however, encounters problems if it used to explain the results of the present experiments, in which the relationship between directness and levels of processing was evaluated by using the *same test cues* at each instructional manipulation. These experiments showed that levels of processing effects were correlated with the directness of the test instructions for fragment, stem, and meaning cues. TAP can explain these dissociations only by assuming that direct instructions involve a greater degree of conceptually-driven processing than indirect instructions for each type of cue. Even if this assumption is granted, this version of TAP would have to incorporate new assumptions to explain how direct instructions are more likely to promote conceptually driven processing for fragment, stem, and meaning cues. An unsolved problem for the approach is to explain how the same cues can be data-driven under some conditions and conceptually driven under others. Some principle in addition to TAP is required.

The TAP framework does not specifically address meaning set size effects—which is unfortunate, because set size can be measured independently of encoding and retrieval tasks and therefore allows one to avoid problems of circularity. Once again, however, if this approach could han-

dle such effects, it would need to incorporate operations that would explain why meaning set size effects linked to the target are obtained with a variety of different types of test cues, but not with others. Such effects are obtained with both meaningfully related cues and weak beginning and ending stem cues, but not with fragments or strongly related stems. In its present form, the TAP framework cannot explain how weak stems, like meaning cues, produce meaning-related target set size effects and how strongly related stems or fragment cues do not produce such effects. Finally, it is unclear how TAP would explain why different aspects of meaning such as levels of processing and meaning set size do not have parallel effects under the same testing conditions. Both manipulations involve meaning, and this version of TAP implies that they should have similar effects, but this parallel was not consistently obtained. For example, the levels of processing manipulation affected fragment completion performance under direct test instructions, but meaning set size had no effect under this condition.

The version of TAP offered by Graf and his colleagues can also explain some but not all of the present findings (Graf & Mandler, 1984; Graf & Ryan, 1990). In this version, direct and indirect tests access two different organizing processes, integration and elaboration. Integration is a result of processing that bonds the features of the target into a unitized representation; elaboration is a result of processing that links the target with other list words, situational information, and *relevant prior knowledge*. Study tasks presumably engage a combination of integrative and elaborative processing and differ in the emphasis placed on each type of processing. Direct and indirect tests tap into elaborative and integrative processing, respectively. This view differs from Roediger's in important details, but it shares the TAP assumption, because performance on both types of tests is presumably determined by study-test overlaps in processing. Indirect tests primarily reflect study-test overlap in integrative processing, and direct tests primarily reflect overlap in elaborative processing.

The TAP framework offered by Graf and his colleagues provides a clear explanation for why levels effects are more apparent under direct than under indirect instructions. Furthermore, because the directness of the test instructions determines what types of study task processing will be recovered at test, this view can explain why this interaction is apparent for meaningfully related cues and for data-driven cues such as word stems and fragments. However, this TAP framework runs into difficulties when applied to meaning set size effects. Although the approach does not speak directly to set size, it links the encoding of relevant prior knowledge to the elaboration process. Meaningfully related associates of the target are presumably incorporated into its elaborative encoding. Given that direct and indirect test instructions respectively tap into elaborative and integrative processes, this view incorrectly predicts that meaning set size effects linked to the target should be more apparent under direct than under indirect

instructions. Such effects, however, were equally apparent under each instructional set. Although the TAP approach offered by Graf and his colleagues can explain effects related to the controlled aspects of meaningful processing, such as manipulations of levels of processing, it does not provide an adequate explanation of effects related to more automatic aspects of processing such as those associated with set size effects.

In contrast to these TAP approaches, PIER was explicitly designed to explain set size effects related to both phonetic and semantic features. It was also designed to deal with the effects of variables influencing the nature of the study trial encoding as well as retention test differences. PIER assumes that different representations are created during the study trial as a result of different processing activities, and that these representations provide mutually exclusive sources of information that can be used separately in a variety of testing situations. Explicit encoding activities create an explicit representation of the target that includes contextual information about the situation and connections to other list items (Anderson & Bower, 1972; Raaijmakers & Shiffrin, 1981). In addition, the implicit activation of the target and its associates creates a second representation. During testing, the explicit and implicit representations can be searched separately or in parallel, depending on the prevailing conditions related chiefly to the nature of the test cue. The success of the explicit search process is a function of how well the target has been encoded during study as determined by intentional encoding operations (e.g., learning strategy, timing and spacing of practice, and so forth) and by the testing conditions (test instructions, presence of interpolated lists, and so forth). The success of the implicit search process depends on how many associates are competing with the target. The smaller the number of competing associates, the more likely it is that the implicit search will be successful (cf. Raaijmakers & Shiffrin, 1981).

PIER attributes the effects of levels of processing and test instructions to the explicit search process. Concreteness ratings produce a more accessible explicit encoding of the target than vowel naming, and, given a related test cue, direct instructions increase the probability that the explicit representation will be searched. The levels × instructions interaction is attributed to the combination of these two factors. Concreteness ratings produce a more accessible explicit encoding than does vowel naming, and this difference should be most apparent when the orientation of subjects is to recall words from the list just presented. Finally, because the search of the explicit encoding can be initiated by any test cue that is related to the target, the levels × instructions interaction should be apparent for a variety of different types of test cues.

PIER attributes the meaning set size effect related to target words to the implicit search process. Because the initiation of this search does not depend on the directness of the test instructions, set size and test instructions are predicted to have additive effects. Furthermore, because explicit and implicit searches produce mutually exclusive contributions to target recovery, the model also predicts

that the magnitude of target set size effects will not be affected by levels of processing during study. Although both levels and set size manipulations in these experiments concern meaning, the model treats these manipulations as tapping into separate aspects of meaning, with levels effects determined by the nature of explicitly controlled encoding activities and set size effects attributed to the automatic activation of related associates. The explicit meaning representation created as a result of concreteness ratings is the result of applying controlled processing operations to a familiar stimulus, whereas the implicit meaning representation created as a result of experiencing the same word is the result of automatic activation processes designed to provide rapid access to related concepts. Either of these representations can lead to target recovery, but, as the model specifies, explicit and implicit searches are exclusively related. The probability of target recovery (P_r) is based on either the success of searching the implicit representation $P(I)$ or on the success of searching the explicit representation, $P(E)$: $P_r = P(I) + P(E)$. Note that, with this equation, the successful recovery of any particular target is based on the success of either search, such that over an entire list the results of these searches will have additive effects on target recovery (see Jones, 1987, for a discussion of exclusion, independence, and inclusion relations between two processes).

Unlike levels effects, meaning set size effects depend on the nature of the test cue. Such effects are found with meaning cues and weakly related stems but not with fragment cues and strong stems. PIER attributes this difference to variations in the nature of the information in the cues. Meaning-related test cues activate their associates and the associates of the target, which then compete for target selection. Weak stems provide partial information about the name or lexical representation for the studied word, but this information is likely to be insufficient for supporting target recovery. This information is supplemented by sampling meaningfully related associates linked to each sampled lexical representation, and the search through these associates is what generates the meaning-related target set size effect for these cues. In contrast with meaning and weak stem cues, strongly related stem cues provide enough information about the name of the target to support its recovery. Strongly related stems have a high initial probability of recovering the name even in the absence of a study trial, and, when combined with a recent exposure, this potential is sufficient to support a high probability of target recovery. As with stem cues, fragment cues provide only partial information, but the test instructions focus subjects on recovering a word name that fits the letter and spacing constraints of the cue, and these restrictions limit the search to lexical information. As a consequence, lexical set size effects are obtained for fragment cues, but meaning set size effects have never been found with such cues (Experiment 1; Nelson, Canas, et al., 1987; Nelson et al., 1989).

By way of final comment, it should be noted that the PIER model has much in common with multiple systems explanations for effects related to implicit memory (e.g.,

Schacter, 1987; Squire, 1987; Tulving, 1985; Tulving & Schacter, 1990). In PIER, searches of explicit and implicit memories are conceptualized as making mutually exclusive contributions to target recovery, and this assumption is compatible with the idea that different memory systems underlie performance in direct and indirect tasks when these tasks are based on the same test cue. Furthermore, set size effects are attributed to the activation of context-free representations, which could be represented as existing in a separate system from representations concerned with more recently acquired episodic information. However, PIER differs from the systems approach in at least two critical ways. First, although compatible with the idea that separate brain systems underlie explicit and implicit memories, PIER is strictly a functional model and makes no assumptions about the intrinsic physical substratum. Second, researchers advocating the systems approach rely on the manipulation of the directness of the test instruction for providing access to different memory systems. Direct instructions presumably provide access to the system underlying the episodic system, whereas indirect instructions increase reliance on systems underlying semantic, procedural, or implicit memory (see Hayman & Tulving, 1989; Richardson-Klavehn & Bjork, 1988; Roediger, 1990; Schacter, 1987, for reviews). In contrast, in the PIER model, test instructions affect the likelihood of searching memories created as a result of explicit encoding operations. The difference is subtle but important. In PIER, test instructions serve only to bias subjects toward engaging in particular types of retrieval operations, just as study instructions (e.g., levels) bias subjects toward engaging in particular types of encoding operations. In this sense, such manipulations are treated as independent variables, not as a means of involving different brain systems.

REFERENCES

ANDERSON, J. R., & BOWER, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, **79**, 97-123.

BLAXTON, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 657-668.

CRAIK, F. I. M., & TULVING, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, **104**, 268-294.

GRAF, P., & MANDLER, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. *Journal of Verbal Learning & Verbal Behavior*, **23**, 553-568.

GRAF, P., & RYAN, L. (1990). Transfer-appropriate processing for implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 978-992.

GRAF, P., & SCHACTER, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 501-518.

HAMANN, S. B. (1990). Level-of-processing effects in conceptually driven implicit tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 970-977.

HASHER, L., & ZACKS, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, **108**, 356-388.

HAYMAN, C. A. G., & TULVING, E. (1989). Contingent dissociation between recognition and fragment completion: The method of triangulation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 228-240.

HOROWITZ, L. M., WHITE, M. A., & ATWOOD, D. W. (1968). Word fragments as aids to recall: The organization of a word. *Journal of Experimental Psychology*, **76**, 219-226.

JONES, G. V. (1987). Independence and exclusivity among psychological processes: Implications for the structure of recall. *Psychological Review*, **94**, 229-235.

LIGHT, L. L., & SINGH, A. (1987). Implicit and explicit memory in young and older adults. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 531-541.

NELSON, D. L. (1981). Many are called but few are chosen: The influence of context on the effects of category size. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 15, pp. 129-162). New York: Academic Press.

NELSON, D. L., BAJO, M. T., & CANAS, J. (1987). Prior knowledge and memory: The episodic encoding of implicitly activated associates and rhymes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 54-63.

NELSON, D. L., CANAS, J., & BAJO, M. T. (1987). The effects of natural category size on memory for episodic encodings. *Memory & Cognition*, **15**, 133-140.

NELSON, D. L., CANAS, J., BAJO, M. T., & KEELEAN, P. (1987). Comparing word fragment completion and cued recall with letter cues. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 542-552.

NELSON, D. L., KEELEAN, P. D., & NEGRAO, M. (1989). Word-fragment cuing: The lexical search hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 388-397.

NELSON, D. L., & MCEVOY, C. L. (1979). Encoding context and set size. *Journal of Experimental Psychology: Human Learning & Memory*, **5**, 292-314.

NELSON, D. L., & MCEVOY, C. L. (1984). Word fragments as retrieval cues: Letter generation or search through nonsemantic memory? *American Journal of Psychology*, **97**, 17-36.

NELSON, D. L., MCEVOY, C. L., JANCZURA, G. A., & XU, J. (1992). Reductions in the influence of implicitly activated memories on recall: Decay, interference, or inhibition. Manuscript submitted for publication.

NELSON, D. L., MCEVOY, C. L., & SCHREIBER, T. A. (1990). Encoding context and retrieval conditions as determinants of the effects of natural category size. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 31-41.

NELSON, D. L., & SCHREIBER, T. A. (1992). Word concreteness and word structure as independent determinants of recall. *Journal of Memory & Language*, **31**, 237-260.

NELSON, D. L., SCHREIBER, T. A., & MCEVOY, C. L. (1992). Processing implicit and explicit representations. *Psychological Review*, **99**, 322-348.

RAAIJMAKERS, J. G. W., & SHIFFRIN, A. M. (1981). Search of associative memory. *Psychological Review*, **88**, 93-134.

RICHARDSON-KLAVEHN, A., & BJORK, R. A. (1988). Measures of memory. *Annual Review of Psychology*, **39**, 475-543.

ROEDIGER, H. L., III (1990). Implicit memory: Retention without remembering. *American Psychologist*, **45**, 1043-1056.

ROEDIGER, H. L., III, & BLAXTON, T. A. (1987). Retrieval modes produce dissociations in memory for surface information. In D. S. Gorfein & R. R. Hoffman (Eds.), *Memory and cognitive processes: The Ebbinghaus Centennial Conference* (pp. 349-379). Hillsdale, NJ: Erlbaum.

ROEDIGER, H. L., III, SRINIVAS, K., & WELDON, M. S. (1989). Dissociations between implicit measures of retention. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit memory: Theoretical issues* (pp. 67-84). Hillsdale, NJ: Erlbaum.

SCHACTER, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 501-518.

SHIFFRIN, R. M., & SCHNEIDER, W. (1977). Controlled and automatic information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127-190.

SQUIRE, L. R. (1987). *Memory and brain.* New York: Oxford University Press.

SQUIRE, L. R., SHIMAMURA, A. P., & GRAF, P. (1985). Independence of recognition memory and priming effects: A neuropsychological analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 11*, 37-44.

SQUIRE, L. R., SHIMAMURA, A. P., & GRAF, P. (1987). Strength and duration of priming effects in normal subjects and amnesic patients. *Neuropsychologia, 25*, 195-210.

SRINIVAS, K. S., & ROEDIGER, H. L., III (1990). Classifying implicit memory tests: Category association and anagram solution. *Journal of Memory & Language, 29*, 389-412.

TULVING, E. (1985). How many memory systems are there? *American Psychologist, 40*, 385-398.

TULVING, E., & SCHACTER, D. L. (1990). Priming and human memory systems. *Science, 247*, 301-306.

TULVING, E., SCHACTER, D. L., & STARK, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 8*, 336-342.

WELDON, M. S. (1991). Mechanisms underlying priming on perceptual tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17*, 526-541.

APPENDIX A

Materials for Experiment 1

	List 1			List 2	
Target	Stem Cue	Fragment Cue	Target	Stem Cue	Fragment Cue
GEM	EM	_EM	FORK	ORK	_ORK
DOG	OG	_OG	GIRL	IRL	_IR_
TUSK	USK	_USK	CALF	ALF	_ALF
LAST	AST	_AST	KEG	EG	_EG
RATTLE	ATTLE	_ATTLE	ROUND	OUND	_OUND
WIFE	IFE	_IFE	THUNDER	UNDER	_ _UNDER
THICK	ICK	_ _ICK	SOIL	OIL	_ _IL
WEST	EST	_E_T	CASH	ASH	_ _SH
DAY	AY	_ _Y	BEE	EE	_ _E
BUMBLE	UMBLE	_ _ _BLE	WIN	IN	_IN
DINE	INE	_INE	MALL	ALL	_ALL
KITE	ITE	_IT_	SAD	AD	_AD
FADE	ADE	_AD_	HAUL	AUL	_ _UL
COLD	OLD	_OLD	CAMP	AMP	_AMP
SAINT	AINT	_AINT	GOAT	OAT	_OAT
THING	ING	_ _ING	YARN	ARN	_ARN
MATCH	ATCH	_ATCH	BREAD	EAD	_ _E_D
ROPE	OPE	_OPE	CARD	ARD	_ARD
DEAL	EAL	_E_L	NECK	ECK	_E_K
SPANK	ANK	_ _ANK	SONG	ONG	_ONG
JOIN	OIN	_OIN	SILK	ILK	_ _LK
RIVER	IVER	_IVER	TELL	ELL	_ELL
WORM	ORM	_ORM	FOOD	OOD	_OOD
GHOST	OST	_ _OST	GLASS	ASS	_ _ASS

Note—The first 12 targets have small meaning set sizes, and the next 12 have large set sizes.

APPENDIX B

Materials for Experiment 2

	Beginning Stems				Ending Stems		
List 1		List 2		List 1		List 2	
Target	Stem Cue	Target	Stem Cue	Target	Stem Cue	Target	Stem Cue
BRAWL	BRA	BRIDE	BRI	BARK	ARK	FAST	AST
CLOSE	CLO	DOG	DO	PONY	ONY	FORK	ORK
NIGHT	NI	MALL	MA	POND	OND	SHOUT	OUT
TALL	TA	THICK	THI	SKETCH	ETCH	THUNDER	UNDER
BAD	BA	NEED	NE	TRIBE	IBE	HIVE	IVE
ZEBRA	ZE	BLADE	BLA	TUSK	USK	LAMP	AMP
BLIND	BLI	CHEAT	CHE	BUMBLE	UMBLE	KEG	EG
CHEESE	CHE	CLOTHES	CLO	CROAK	OAK	LIMB	IMB
FLAG	FLA	FRESH	FRE	DAY	AY	WREN	EN

APPENDIX B (Continued)

Beginning Stems				Ending Stems			
List 1		List 2		List 1		List 2	
Target	Stem Cue	Target	Stem Cue	Target	Stem Cue	Target	Stem Cue
FROG	FRO	STICK	STI	MINK	INK	WIFE	IFE
STICK	STI	TRUCK	TRU	SABER	ABER	NOTION	OTION
SOFT	SO	SOFT	SO	SLIM	IM	RATTLE	ATTLE
BLAZE	BLA	BROTH	BRO	BABY	ABY	FOOD	OOD
FINISH	FI	CROWN	CRO	BEG	EG	HANG	ANG
HAPPY	HA	FRAME	FRA	CAR	AR	HOPE	OPE
NEEDLE	NE	FROST	FRO	FISH	ISH	MESS	ESS
SALT	SA	QUIZ	QUI	SCHOOL	OOL	PLEASE	EASE
THIN	THI	THUNDER	THU	FOG	OG	SILK	ILK
BOTTLE	BO	BASKET	BA	CARD	ARD	CASE	ASE
CRAB	CRA	CRUTCH	CRU	DAMP	AMP	DIRT	IRT
DANCE	DA	DESK	DE	GET	ET	FADE	ADE
GRAVE	GRA	GRACE	GRA	STAGE	AGE	SLAVE	AVE
WOOL	WO	NUMBER	NU	THING	ING	SPELL	ELL
MOOSE	MO	SHAKE	SHA	WEAVE	EAVE	JOIN	OIN

Note—The first six targets in each list have small meaning sets and strongly related stem cues, the next six have large sets and strongly related stem cues, the next six have small meaning sets and weakly related stem cues, and the final six have large meaning sets and weak stem cues.

APPENDIX C

Materials for Experiment 3

Beginning Stems				Ending Stems			
List 1		List 2		List 1		List 2	
Target	Stem Cue	Target	Stem Cue	Target	Stem Cue	Target	Stem Cue
CALF	CA	DUMB	DU	BROTH	OTH	CAT	AT
DIM	DI	BRIDE	BRI	CLOCK	OCK	BRUSH	USH
FRAME	FRA	BROTH	BRO	CROWN	OWN	CROAK	OAK
FRIGID	FRI	CAT	CA	CUP	UP	FAST	AST
HOT	HO	DIG	DI	FRAME	AME	FLESH	ESH
KNOB	KNO	FLESH	FLE	GLOBE	OBE	FROST	OST
SOUTH	SOU	FROST	FRO	HIVE	IVE	BARK	ARK
SPEND	SPE	HAUNT	HAU	KNOB	OB	STOP	OP
SWIFT	SWI	NAIL	NA	MUD	UD	SAD	AD
THICK	THI	QUIZ	QUI	SKUNK	UNK	SHOUT	OUT
THUMB	THU	THIN	THI	SPEND	END	STAY	AY
WHITE	WHI	RAKE	RA	SWIFT	IFT	TRIBE	IBE
BLANK	BLA	BLOOD	BLOO	BRIBE	IBE	BEG	EG
BRIBE	BRI	COLD	CO	BUS	US	BUG	UG
BUG	BU	CRAB	CRA	CRAB	AB	CRISP	ISP
CLEAR	CLE	CRUNCH	CRU	DRUM	UM	FLAG	AG
COAT	CO	BAG	BA	FLASK	ASK	GRAPE	APE
CRUTCH	CRU	FADE	FA	FRESH	ESH	HANG	ANG
FACE	FA	FINE	FI	GRAVE	AVE	HEAT	EAT
FLASK	FLA	GRAPE	GRA	GROUP	OUP	HIP	IP
GUN	GU	LEAVE	LEA	MESS	ESS	SWIM	IM
GREEN	GRE	MAID	MAI	PLAN	AN	TAX	AX
GROUP	GROU	STUFF	STU	RAFT	AFT	THING	ING
SHAPE	SHA	SING	SI	SHAPE	APE	TRUCK	UCK

Note—The first 12 targets in each list have small meaning sets, and the next 12 have large meaning sets.