ACADEMIC PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Memory and Language 48 (2003) 563–580

Journal of Memory and Language

www.elsevier.com/locate/jml

Remembering and knowing in context

Glen E. Bodner[a,*] and D. Stephen Lindsay[b]

[a] *Department of Psychology, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada T2N 1N4*
[b] *Department of Psychology, University of Victoria, P.O. Box 3050 STN CSC, Victoria, BC, Canada V8W 3P5*

Received 22 April 2002; revision received 1 August 2002

Abstract

Prior to a recognition test, subjects studied one set of words in a medium level of processing (LOP) task and another set of words in either a shallow or deep LOP task. Medium items received more remember judgments (and fewer know judgments) when mixed with shallow than with deep items (Experiment 1)—even when a basis was required for each remember judgment (Experiment 4). These effects were due to the test-list context: judgments for medium items were equivalent for the two groups when only the medium items were presented at test (Experiment 2). The relative weighting subjects assigned to particular kinds of recollected information as the basis of their remember judgments was affected by list context (Experiment 4), but their ability to remember list source was not (Experiment 3). The test-list context appears to have influenced subjects' functional definitions of remembering and knowing rather than the contents of their recollections.

Keywords: Remember/know judgments; List context effects; Recognition memory; Source memory

Several theories of recognition memory distinguish between experiences of "remembering," which are accompanied by recollection of episodic details from an earlier encoding experience, and experiences of "knowing," which are not (Gardiner & Java, 1993; Jacoby, 1991; Mandler, 1980; Rajaram, 1993; Tulving, 1985; Yonelinas, 2001). Mandler (1980) provided the classic example of knowing without remembering: when seeing a familiar person in an unusual context (e.g., the butcher on the bus) gives rise to a maddening feeling of *knowing* one has seen the person before without being able to *remember* who the person is. Here, we examine the possibility that whether people will report remembering versus knowing that they have studied an item will depend on the quality of their memories for another set of items on the test list.

* Corresponding author. Fax: +403-282-8249.
E-mail address: bodner@ucalgary.ca (G.E. Bodner).

Theoretical perspectives on the remember/know distinction

The distinction between the phenomenological experiences of remembering and knowing has been extensively researched by John Gardiner and his colleagues (for recent summaries, see Gardiner & Conway, 1999; Gardiner & Richardson-Klavehn, 2000). They and others have reported numerous studies demonstrating that judgments of remembering and knowing can be functionally independent. Some experimental manipulations selectively affect only one type of judgment whereas others have opposite or parallel effects on the two judgments.

To take an example relevant to our experiments, Gardiner (1988) had subjects study a list of words in a task designed to induce either a relatively shallow level of processing (producing a rhyme for each word) or a relatively deep level of processing (producing a semantic associate for each word). At test, subjects in both groups were given a recognition test consisting of studied and

0749-596X/02/$ - see front matter
doi:10.1016/S0749-596X(02)00502-8

new words. They circled the words they recognized, and reported whether they recognized each circled word in the remember or know sense. Deeper processing at encoding led more to remember responses than shallower processing (e.g., Craik & Tulving, 1975), but this level of processing (LOP) manipulation had no effect on the rate of know responses (but see Rajaram, 1993).

Dissociations such as the effect of LOP on remembering but not on knowing suggest that these two states of awareness may be phenomenological markers of different memory systems or processes. Indeed, prior to many of these demonstrations, Tulving (1985) proposed that remembering reflects a state of autonoetic consciousness produced by an episodic memory system, whereas knowing reflects a state of noetic consciousness produced by a semantic memory system. Similarly, Gardiner (1988) suggested that remember judgments measure explicit memory, whereas know judgments measure implicit memory. Remembering has also been described as an outcome of prior conceptual processing, and knowing as an outcome of prior perceptual processing (Gardiner & Java, 1993; Rajaram, 1993). The influence of LOP on remembering could thus reflect enhanced activation in an episodic memory system and/or greater conceptual processing for deeper items relative to shallower items. In turn, the finding that LOP does not influence knowing could reflect equivalent semantic system activation and/or equivalent perceptual encoding for deeper and shallower items. More recently, Rajaram (1996, 1998) has developed a more flexible account in which remembering increases when the distinctiveness or salience of items is maximized, and knowing increases when processing fluency is maximized.

A dual-process model of recognition memory (e.g., Jacoby, 1991; Jacoby & Dallas, 1981) has also been used to understand data from the remember/know paradigm. This model postulates that people have two independent bases available for making recognition judgments: a controlled recollection process and an automatic familiarity process (see also Mandler, 1980). Yonelinas and Jacoby and their colleagues (Jacoby, 1998; Jacoby, Yonelinas, & Jennings, 1997; Yonelinas & Jacoby, 1995) have suggested that remember judgments can provide a good estimate of a threshold recollection process, R, and that the proportion of non-recollected items that receive know judgments, know/(1-R), provides an estimate of a signal-detection familiarity process, F. Thus, although remember/know reports are mutually exclusive at the level of responses (cf. Higham & Vokey, 2002), it has been argued that the processes that give rise to these reports are independent.

The necessity of positing that remembering and knowing reflect the operation of distinct systems or processes has been questioned by researchers advocating a two-criterion signal-detection model (Donaldson, 1996; Hirshman & Master, 1997; Inoue & Bellezza, 1998). These researchers model recognition memory decisions using a single-process signal-detection model in which studied items form a distribution of memory strength. Subjects set a yes/no criterion along this continuum, and will only claim to recognize items falling above this criterion. To model remember/know judgments, it is assumed that subjects set a stricter criterion to divide recognized items into those that are remembered and those that are known. The criterion model accommodates dissociations in the effects of variables on remembering versus knowing by assuming that subjects shift their decision criteria depending on the experimental conditions (see Donaldson, 1996, or Hirshman & Master, 1997, for examples). The criterion model explanation of remember/know judgments has generated a fair amount of controversy (e.g., Gardiner & Conway, 1999; Gardiner & Gregg, 1997; Gardiner, Richardson-Klavehn, & Ramponi, 1998; Hirshman, 1998), in large part because differential placement of decision criteria does not explain why remembering and knowing are consciously experienced so differently.

Gruppuso, Lindsay, and Kelley (1997) proposed an account of the remember/know distinction that combines aspects of dual-process models and single-process criterion models. Following global models of recognition memory (e.g., Gillund & Shiffrin, 1984; Hintzman, 1986; Murdock, 1982) and the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993), Gruppuso et al. assumed that a studied item presented on a test list will evoke a subset of the multiple features or aspects of the encounter with that item on the study list. They further assumed that whether the evoked information gives rise to an experience of remembering or an experience of knowing depends not on the quantity of this information, but rather on whether it enables the person to perform the memory task at hand (e.g., identifying the person on the bus as the butcher).

Gruppuso et al. (1997) reported a series of studies using an adaptation of Jacoby's (1991) process-dissociation procedure. This procedure yields separate estimates of the contribution of F and R to responding on a recognition test. Subjects study words from two lists and are later shown words from both lists mixed with new words. These types of items are presented in two conditions. In the inclusion condition, subjects are to say "yes" to items from either list and "no" to new items, whereas in the exclusion condition they are to say "yes" only to items from list 1, and "no" to studied items from list 2 and to new items. The estimate of R is computed as the difference between the rate of yes responses to list 2 items in the inclusion versus exclusion conditions, and the estimate of F is based on the rate of yes responses to non-recollected list 2 items.

Using this procedure, Gruppuso et al. showed that whether memory for the encoding task used for a given item contributed to the estimate of R or to the estimate of F depended on whether it enabled the subject to identify the list in which the item had been studied. When the same task was used to encode items on both lists, then memory of the encoding task contributed to the estimate of F because although it provided a basis for recognizing the item, it did not provide a basis for excluding the item as a list 2 item. When different (but equally "deep") encoding tasks were used for each list, then memory for the encoding task contributed to the estimate of R because it enabled subjects to reject items from the to-be-excluded list (list 2).

Other investigators (Dodson & Johnson, 1996; Mulligan & Hirshman, 1997) interpreted similar findings as evidencing flaws in the process-dissociation procedure, but Gruppuso et al. instead proposed that recollection and familiarity should be defined in functional terms: recollection occurs when memory information enables one to perform the task at hand (e.g., identifying the familiar person on the bus as your butcher) whereas familiarity occurs when memory information enables one to recognize a stimulus as familiar but does not enable one to perform the task at hand (e.g., being unable to identify the familiar person on the bus). According to Gruppuso et al., some kinds of memory information typically contribute to familiarity rather than to recollection (e.g., memory for low-level, data-driven, automatic processes such as those involved in perceptually identifying a word) because they typically are not uniquely source specifying. Other kinds of memory information typically contribute to recollection rather than familiarity (e.g., memory for consciously controlled reflective processes) because they often uniquely specify a particular episode from one's past. Importantly, however, whether a particular piece of information contributes to familiarity or recollection depends on the specifics of the study and test situations (see also Conway & Dewhurst, 1995a; Dewhurst & Conway, 1994; Whittlesea, 2002b; Yonelinas & Jacoby, 1996).

Context and remember/know judgments

Our goal in the present experiments was to examine how one specific aspect of the test situation—the memorability of the test-list context items—influences the recognition judgments made for a critical set of items. To this end, we devised a paradigm that employs three distinct LOP tasks, which we label shallow, medium, and deep. All subjects studied one set of words using the medium LOP task, and studied a second set of words using either the shallow or deep LOP task. Thus, there were two groups in each experiment: *medium-with-shallow* and *medium-with-deep*. After both lists were presented, subjects were given a remember/know recognition test consisting of studied words from both sets plus an equal number of new words. The manipulation of LOP at encoding was expected to have a large influence on recognition performance, of course, but we were chiefly interested in examining how remember/know judgments (Experiments 1and 4) and list source judgments (Experiment 3) for the medium LOP items were influenced by the inclusion on the test list of a second set of items that had been studied in either a relatively more memorable (deep LOP task) or less memorable (shallow LOP task) way. In Experiment 2, we examined whether having studied two lists of words would produce a list-context effect at test when in some conditions only the medium LOP items were included on the test list.

Our paradigm is similar to those used to investigate the list-strength effect, in which the "strength" with which items are encoded is manipulated at study (typically by increasing the number or the duration of exposures to a subset of the items). At test, memory for items in pure lists (those containing either all strong or all weak items) is compared to memory for items in mixed lists (those containing a mixture of strong and weak items). Recall of strong items is often better in mixed lists than in pure-strong lists, presumably because strong items are more distinctive when in the context of weak items, and recall of weak items is better in pure-weak lists than in mixed lists, presumably because memory for the strong items in mixed lists interferes with retrieval of the weak items (see Shiffrin, Ratcliff, & Clark, 1990).

Although list strength has been found to affect recall, it has typically been found not to influence yes/no recognition judgments (e.g., Hirshman, 1995; Murnane & Shiffrin, 1991a, 1991b; Ratcliff, Clark, & Shiffrin, 1990; Ratcliff, Sheu, & Gronlund, 1992; Yonelinas, Hockley, & Murdock, 1992). Our experiments explore the possibility that although list strength (or list context more generally) may not typically affect whether an item is recognized, it may affect whether the item is recognized in the remember sense or in the know sense (e.g., Macken, 2002). If the states of awareness that accompany recognition are influenced by list context, then manipulations of list context become useful tools for investigating how context contributes to the construction of subjective experience (e.g., Dewhurst & Parry, 2000) and for constraining theoretical accounts of what remember/know judgments represent.

Experiment 1

Method

Subjects. Subjects in Experiment 1A were University of Victoria undergraduates. Subjects in Experiment 1B and in all remaining experiments were University of Calgary undergraduates. All subjects volunteered for extra course credit. The 48 subjects in each experiment were randomly assigned to either the medium-with-shallow or the medium-with-deep group ($n = 24$ per group). To ensure that all subjects closely adhered to the remember/know instructions, two subjects in each experiment who falsely remembered more than 10% of the new words were replaced.

Materials and design. The stimuli were 128 five-letter words between 0 and 680 in frequency ($M = 50$, $SD = 94$; Kučera & Francis, 1967) that ranged from concrete (e.g., salad, beach, wrist) to abstract (e.g., while, fraud, glory) in nature. The 120 critical words were divided into four blocks of 30, two of which served as new words, one as the deeper LOP list, and one as the shallower LOP list, according to a counterbalanced schedule. The remaining eight words served as buffers; two were presented at the beginning and two were presented at the end of each study list. The order of the two study lists was counterbalanced (e.g., half of each group studied the shallower list first and half studied the deeper list first), as was the order of the words on the study lists and test list.

Procedure. Individually tested subjects were told that they would be studying two lists of words under different instructions for a later memory test. The instructions for the first list were given, then the list was presented one word at a time on a computer monitor. This procedure was repeated for the second list. Each task involved making a verbal yes/no judgment about each word on the list: whether or not the word contained the letter A (shallow LOP task), whether or not the word is one people commonly use (medium LOP task), or whether or not the word refers to something the subject would want to have if he/she were stranded on a desert island (deep LOP task). We were not interested in LOP per se, but rather in how the subjective experience of recognition might be influenced by the memorability of the other items on the test list. Therefore, to ensure that we would have three distinct levels of performance, shallow items were displayed for 500 ms, medium items for 1 s, and deep items for 2 s, thus confounding presentation duration with the LOP task. In addition, subjects read medium and deep items aloud before making judgments for these items.

After the two lists, the recognition instructions were presented. Subjects were told that they would be seeing a list of words on the computer screen one at a time, some of which had been presented during the study phase, and that they were to classify their recognition experience for each word into one of three distinct categories. They were told to say *remember* only if they were sure that they recognized the word and *only* if they also remembered some specific detail about their study experience with it (e.g., a specific thought, image, or feeling they experienced at the time they studied it). They were told to say *know* if they were *sure* they recognized the word but could not remember any specific details about their study experience with it. They were told to say *neither* if they did not think they had studied the word or if they were unsure whether they had studied the word or not. Before beginning the test, subjects were asked to summarize, in their own words, the difference between remembering and knowing. The recognition test was initiated only when the experimenter was satisfied that the subject fully understood the instructions. To further encourage careful responding, subjects made their recognition judgments aloud, and the experimenter recorded their responses. After the 120-item recognition test, subjects were asked to provide examples of words that they had remembered, and were asked to explain how they had distinguished remembering from knowing. Each testing session took 20–25 min.

Experiments 1A and B were essentially replications at different universities, with different experimenters. One other difference was that in Experiment 1A, presentation of each study list item was contingent on the experimenter keying in the subject's response to the previous item, whereas in Experiment 1B, the computer presented list words at a 750-ms inter-stimulus interval and the experimenter wrote subjects' responses on a sheet of paper. This change in Experiment 1B was made to ensure that the pacing of the study lists was the same for all subjects.

Results and discussion

Fig. 1 shows the proportion of remember and know judgments for each set of items on the recognition test for the medium-with-shallow and medium-with-deep groups in Experiments 1A and B. The sum of remember and know judgments represents overall recognition. Although different experimenters collected the data at different universities, the results of Experiments 1A and B were nearly identical. Experiment was included as a variable in exploratory analyses, and produced no reliable effects. In describing the results below, we therefore collapse the data from the two experiments. The effects we report were significant at the $p < .05$ level unless otherwise noted.

Task order. The order of the two LOP tasks at study (deeper task vs. shallower task first) was included as a factor in each analysis of variance (ANOVA) reported in this paper. Importantly, task order did not interact with any of the list context effects we report, with one exception (see Experiment 3). Thus, in general, how subjects encoded their first list of words did not have a significant influence on how they encoded their second list of words (e.g., there were no carryover effects).

Appropriate use of judgments. After the experiment, all subjects were able to provide examples of items

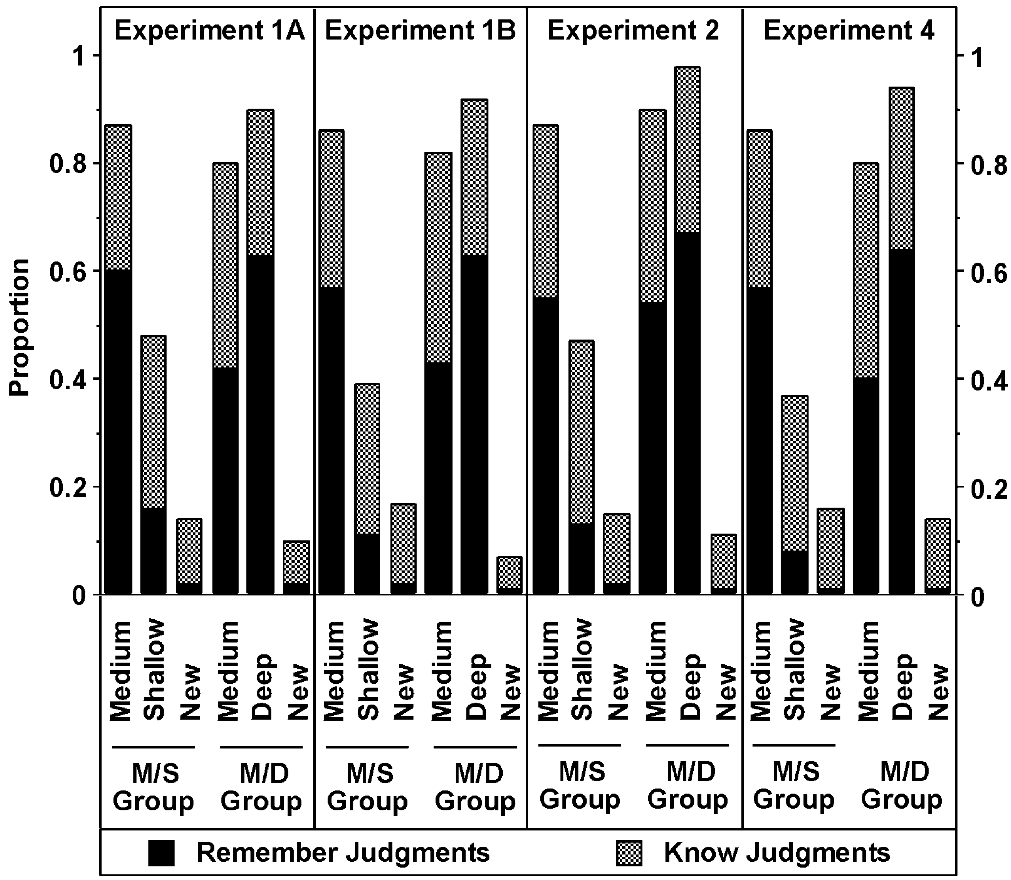

Fig. 1. Mean proportion of remember and know judgments for studied and new words in the medium-with-shallow (M/S) and medium-with-deep (M/D) groups in Experiments 1A and B, 2, and 4. The sum of remember and know judgments estimates overall recognition.

classified as remember and to state why they remembered these items. Subjects were not typically able to provide an example of an item they had classified as know, but they all appropriately explained what led them to make know judgments. The low levels of remember ($M = .02$, $SD = .02$) and know ($M = .10$, $SD = .10$) responses to new words support the experimenters' impressions that subjects used the remember/know categories appropriately. The medium-with-deep group, who were given the most memorable pair of encoding tasks, made fewer know judgments to new items (.07 vs. .13), $F(1,92) = 12.56$, $MSE = .01$, $\eta^2 = .12$, and nonsignificantly fewer remember judgments to new items (.01 vs. .02), $F(1,92) = 1.50$, $MSE = .0004$, $p = .22$, $\eta^2 = .02$, than the medium-with-shallow group.

Level of processing effects. As anticipated, our encoding manipulations were effective in producing three distinct levels of recognition as measured by overall recognition and by remember judgments (shallow < medium, medium < deep; all ps < .0001). In the medium-with-shallow group, know judgments were equally likely for medium and shallow items (.28 vs. .30), $F < 1$. In the medium-with-deep group, know judgments were more likely for medium items than deep items (.38 vs. .28), $F(1,46) = 26.63$, $MSE = .01$, $\eta^2 = .37$. Thus, in one case, the LOP/duration manipulation did not affect know judgments, as Gardiner (1988) found, and in the other case an increase in LOP was associated with lower rates of knowing, as Rajaram (1993) reported. Gardiner, Java, and Richardson-Klavehn (1996) suggested that reverse LOP effects on knowing occur when subjects include guesses in the know category. While we cannot rule out this possibility, the rate of know judgments was the same for each set of items in the medium-with-shallow group, and our subjects were certainly not liberal in their use of know judgments, as indicated by their low rates of know judgments to new words.

In contrast to the variable effects of LOP/duration on know judgments, dual-process estimates of familiarity (excluding estimates of 0 and 1) were greater for the medium items than for the shallow items in the medium-with-shallow group (.65 vs. .35), $F(1,39) = 86.46$,

$MSE = .02$, $\eta^2 = .69$, and were near-significantly greater for the deep items than for the medium items in the medium-with-deep group (.70 vs. .65), $F(1, 34) = 3.98$, $MSE = .01$, $p = .054$, $\eta^2 = .10$.[1] This pattern is consistent with 16 out of 17 studies analyzed by Yonelinas (2002) and suggests that deeper LOP increases familiarity on tests of recognition memory (see also Jacoby & Kelley, 1991).

List-context effects. As can be seen in Fig. 1, the overall level of recognition of medium items was not greatly influenced by whether they were mixed with shallow or deep items (i.e., by the test-list context). However, the small advantage for recognition of medium items in the medium-with-shallow group over the medium-with-deep group did brush significance (.86 vs. .81), $F(1, 92) = 3.89$, $MSE = .02$, $p = .052$, $\eta^2 = .04$. Most importantly, in contrast to the small effect on overall recognition, list context substantially affected whether subjects classified their recognition of medium items as an experience of remembering or of knowing. Medium items were 15% more likely to be classified as remember when mixed with shallow items rather than deep items (.58 vs. .43), $F(1, 92) = 15.04$, $MSE = .04$, $\eta^2 = .14$, and were 10% less likely to be classified as know (.28 vs. .38), $F(1, 92) = 10.41$, $MSE = .02$, $\eta^2 = .10$.

Another way of assessing the influence of list context is to examine the proportion of recognized medium LOP items that were classified as remember (i.e., the remember:recognize ratio), since there is no necessary relationship between making a recognition judgment and making a remember judgment (Rajaram, 1993). In the medium-with-shallow group, .67 of the medium items that were recognized were classified as remember, whereas this ratio was only .52 in the medium-with-deep group. The difference between these two ratios was significant, $F(1, 92) = 15.32$, $MSE = .04$, $\eta^2 = .14$, again showing that list context had a robust effect on how people classified their recognition experiences.

Dual-process theories claim that a remember judgment will be made only if the threshold recollection process succeeds at retrieving some information about the study event (e.g., Yonelinas, 2001). To accommodate our results, the dual-process account would have to claim that recollection of medium items was less likely to succeed if the other set of test items had been studied at a deeper level rather than at a shallower level. The nearly significant effect of list context on overall recognition is consistent with such a claim. However, there was no influence of list context on estimates of the familiarity process for medium items (.65 vs. .64), $F < 1$ (excluding estimates of 0 and 1, see Footnote 1), which is the process that should have been the most affected by list context if list context influences response criteria (Yonelinas, 2002).

Experiment 1 clearly shows that the test-list context (in addition to encoding factors such as LOP) influences whether subjects will claim to remember or know that they have studied a set of items. Remember judgments for medium LOP items were more likely (and know judgments were less likely) when these items were presented in a list containing a set of shallow LOP items as opposed to a set of deep LOP items. These list-context effects were induced quite incidentally by the nature of the recognition context, rather than being produced by an instructional manipulation of subjects' recognition decision criteria (cf. Hirshman & Henzler, 1998; Strack & Förster, 1995). They were not the result of sloppy use of the remember and know categories; remember and know judgments to new words were relatively rare and subjects defined remembering and knowing appropriately both before and after the experiment.

Experiment 2

The order of the two LOP tasks did not interact with the list-context effects observed in Experiment 1, suggesting that the nature of the first LOP task did not influence how subjects performed their second LOP task. This implies that remember/know judgments for medium LOP items were instead influenced by the inclusion on the test list of another set of items that systematically differed in memorability. Experiment 2 was designed as a further test of the hypothesis that the effect of list context arises during test rather than as an artifact of differential processing at study. In essence, Experiment 2 was a replication of Experiment 1B with the crucial exception that only one of the two sets of LOP items was presented along with the set of new items on the test list. The critical conditions are those in which the medium items were included on the test list. If the nature of one encoding task influences how the other encoding task is performed or how recognition experiences are classified at test, then the list-context effects found for medium items in Experiment 1 should be replicated in Experiment 2. If list-context effects instead arise during the test as a consequence of making judgments about items studied at different levels of processing, then no list-context effects for medium LOP items should be obtained in Experiment 2.

Method

Subjects. Twenty-four subjects were randomly assigned to each group, labeled as before. We replaced two

[1] Familiarity estimates at floor or ceiling (i.e., 0 or 1) were excluded from the analysis. In the medium-with-shallow group, six familiarity estimates for medium items were at ceiling, and one estimate for shallow items was at floor. In the medium-with-deep group, two familiarity estimates for deep items were at floor and nine were at ceiling, and one estimate for medium items was at ceiling.

subjects who falsely remembered more than 10% of the new words.

Materials, design, and procedure. Experiment 2 was the same as Experiment 1B, except that only one of the two LOP lists was included on the recognition test, randomly mixed with 30 new words. Ninety of the 120 words from Experiment 1 were randomly selected for each subject. Thirty were assigned to each of the two study lists and 30 served as the new words.

Results and discussion

Subjects defined and used the remember and know categories appropriately; as a result, remember ($M =$ .02, $SD =$ .03) and know ($M =$ .11, $SD =$ .10) judgments to new words (averaged across subjects tested on each list within each group) were infrequent. As Fig. 1 shows, the LOP encoding manipulations again produced three distinct levels of memory performance. The overall level of recognition of medium items was similar to that found in Experiment 1, and was not influenced by whether the other LOP task provided at study was the shallow or deep LOP task (.87 vs. .89), $F < 1$. Most importantly, whether the shallow or deep LOP task had been used to study the non-tested second set of items had no influence on the rate of remember judgments (.55 vs. .54) or know judgments (.32 vs. .36) given to medium items, Fs < 1.

The absence of list-context effects in Experiment 2, along with the absence of task order effects in Experiments 1 and 2, suggest that the list-context effects in Experiment 1 were not the result of differential encoding of items on the second encoding task due to carrying over the LOP instructions used for the first encoding task. Subjects in Experiment 2 were as likely to make remember and know judgments to medium items whether the other set of items had been studied in a deep or shallow LOP task. The absence of list-context effects in Experiment 2 is also inconsistent with the possibility that our subjects set their remember/know criteria immediately before the test based on the products of a spontaneous attempt to free-recall items from both lists (see Hirshman, 1995). Although remember/know criteria may sometimes be set in this manner, had our subjects done so, list-context effects should have been obtained in Experiment 2. Instead, we suggest that when only the medium LOP items were included on the test list, subjects in both list context groups had similar kinds of recognition experiences at test, and hence both defined remembering and knowing similarly.

Dewhurst and Parry (2000) reported a list context effect in which emotional words were more likely than neutral words to receive remember judgments when both sets were mixed together at both study and test, but remember judgments were equally likely for both types of words when pure blocks of emotional and neutral words were presented at both study and test. Although it is unclear from their design whether list context affected encoding or recollection processes, the same conclusion reached by Dewhurst and Parry is bolstered by the results of Experiments 1 and 2, namely that attributions of remembering can be influenced by the structure of the recognition test list.

Experiment 3

As noted above, the lack of reliable list-context effects on overall recognition (Experiments 1 and 2) and the absence of carryover effects of encoding instructions on remembering and knowing when the test-list context was kept constant between groups (Experiment 2), are inconsistent with the idea that list context changes the information subjects recollect about test items. Instead, we suggest that during the recognition test, subjects learn to "tune into" those aspects of their recollections that they feel are the most effective ones for making their recognition judgments. Subjects in different list-context groups may therefore be equally able to recollect a particular attribute about their study experiences, but based on their different list-context-driven experiences during the test list, they may come to emphasize different recollected attributes when making judgments about their subjective recognition states.

To further test this claim, in Experiment 3 we replaced the remember/know test with a list source judgment test. Following the same encoding phases in Experiments 1B and 2, subjects classified each word on the test list into one of four categories: as a word they recollected was on list 1, as a word they recollected was on list 2, as a word they were sure was studied but were unsure of on which list, or as a new word. If list context affects what subjects *are able to recollect* about their study experiences, then the medium-with-shallow group, who reported more remembering in Experiment 1, should be more likely than the medium-with-deep group to recollect list source for medium LOP items, even when both groups are instructed to focus on recollecting list source information. In contrast, if list context affects what recollected information subjects *tend to use*, then the two groups might be equally able to recollect list-source-specifying information for medium LOP items when the task requires that such information be accessed.

Prior research on source monitoring indicates that, all else being equal, source monitoring accuracy decreases as memories from the potential sources become more similar to one another (Johnson et al., 1993). Therefore, assuming that memories of shallow and medium items differ from one other to a similar extent as memories of medium and deep items, source monitoring for medium items should be roughly equivalent in the two groups. Consequently, equal source monitoring in the two groups could be taken as evidence that the list-context effects on remembering and knowing in Experiment 1 were driven by differences in the relative weighting the two groups assigned to different recollected attributes, rather than to differences in the amount of information that each group could potentially recollect.

Method

Subjects. Twenty-four subjects were randomly assigned to each group. We replaced six subjects (three from each group) who incorrectly judged more than 10% of the new words to have been presented on the medium or deep LOP list (it proved difficult to prevent the medium-with-shallow group from attributing more than 10% of the new words to the shallow LOP list).

Materials, design, and procedure. Experiment 3 was the same as Experiment 1B except that at test subjects were given a list-source task rather than a remember/know task. Subjects were to say "list 1" or "list 2" *only* if they were sure that they recognized the word and *only* if they recollected that the word had appeared on that list. They were then reminded of the LOP task they had performed for each list. Subjects were instructed to say "unsure" if they were not sure which list the word was on but were confident that the word had appeared on one of the two lists. Finally, they were asked to say "new" when they thought the word was a word that they had not studied or when they were unsure whether the word had been studied or not. Subjects summarized the instructions and the experimenter clarified any misunderstandings before the test began.

Results and discussion

The left panel of Fig. 2 shows the proportion of correct, incorrect, and unsure source judgments for each set of studied words in the two groups. The level of overall recognition of medium items (sum of correct, incorrect, and unsure judgments) was not reliably different for the medium-with-shallow and medium-with-deep groups (.89 vs. .87), $F < 1$.

Source attributions for studied words. Of greatest interest here is the finding that the proportion of correct source judgments for medium items was not different in the medium-with-shallow and medium-with-deep groups (.60 vs. .57), $F < 1$. This null effect was not due to a lack of sensitivity, because the accuracy of source judgments was robustly influenced by LOP (shallow < medium, medium < deep; both ps $< .0001$). Moreover, the power to detect an effect of list context on correct source judgments of the same magnitude as the effect of list context on remember judgments in Experiment 1 was a respectable .78 ($d = .81$). Unsure judgments to medium items were also not influenced by list context, $F < 1$.

We take the equivalence of correct list-source judgments in the two groups as evidence that the list-context manipulation did not affect subjects' ability to recollect information about studied words. The results of

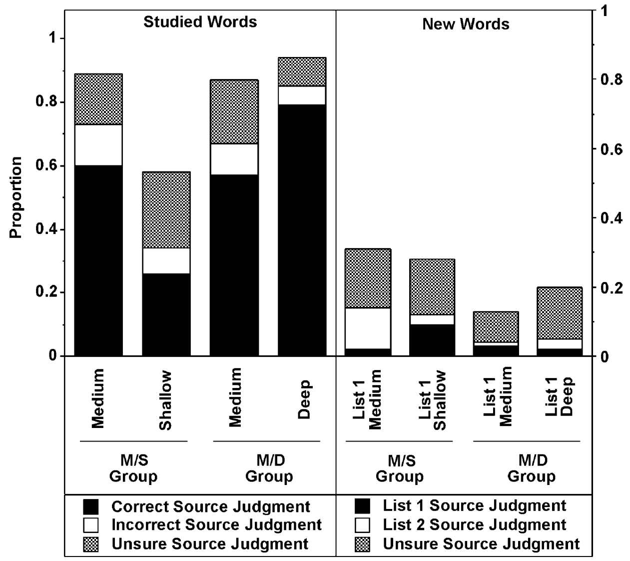

Fig. 2. Mean proportion of correct, incorrect, and unsure source judgments for studied words in the medium-with-shallow (M/S) and medium-with-deep (M/D) groups (left panel), and mean proportion of List 1, List 2, and unsure source judgments for new words as a function of list context and which of the two lists was List 1 in Experiment 3 (right panel). The sum of correct, incorrect, and unsure judgments estimates overall recognition.

Experiment 3 thus suggest that list context in Experiment 1 affected subjects' definitions of what constitutes remembering versus knowing, rather than the contents of their recollections and/or their ability to recollect particular details about their study experiences. Of course, our findings demonstrate this equivalence only for recollection of details that allow list source to be specified, and it is possible that the list-context manipulation affected retrieval of other kinds of non-source-specifying information. Nonetheless, we take the null effect in Experiment 3 as evidence that the test-list context mediated subjects' notions of what constitutes remembering rather than changing the amount of information subjects retrieved at test.

At first glance, our findings that the list-context manipulation affects remember/know judgments but does not affect source judgments appear to conflict with prior evidence of convergence between these judgments. Conway and Dewhurst (1995a) found similar patterns of data whether subjects made remember/know or source judgments (see also, Donaldson, MacKenzie, & Underhill, 1996), leading them to suggest that the state of awareness that accompanies memory for source attributes also accompanies memory for other recollected attributes. In addition, Rugg, Schloerscheidt, and Mark (1998) found that event-related brain potentials associated with remember/know and source judgments were equivalent, suggesting that the recollection process was the same in the two tasks. However, the two kinds of judgments will not always be isomorphic, as the following thought experiment demonstrates. Imagine a two-list encoding phase consisting of two highly similar encoding tasks that both encourage subjects to develop idiosyncratic thoughts (e.g., task 1, produce a free association; task 2, produce a highly related word). Each encoding task should induce a relatively deep LOP, and hence should produce a relatively high rate of remember judgments. But because each task will yield recollections that are not particularly diagnostic of source, the rate of correct source judgments should be much lower than the rate of remember judgments to studied items (for a related demonstration, see Lindsay & Johnson, 1991).

Although source judgments were not influenced by test-list context as manipulated in Experiment 3, source judgments would very likely be influenced by other manipulations of test-list context. For example, one might be able to design medium and deep LOP tasks that are more distinctive from one another than are the medium and shallow LOP tasks, which would promote better source judgment accuracy for medium items in the medium-with-deep group. Regardless, how context shapes recollective experience appears to be a tractable issue, and an important question worthy of further attention is whether context and/or task variables might affect subjects' ability to recollect attributes other than those which allow list membership to be specified.

Source attributions for new words. The right panel of Fig. 2 displays the proportion of false source attributions for new words (saying "list 1," "list 2," or "unsure" to a new word) as a function of the order of the two lists for each group.[2] For example, "List 1 Shallow" refers to the medium-with-shallow subjects for whom list 1 was shallow LOP and list 2 was medium LOP, and "List 1 Medium" refers to the medium-with-shallow subjects for whom this order was reversed. Interestingly, subjects showed an "it-had-to-be-shallow" bias in which they were more likely to claim that a new word had been presented on whichever was the shallower of the two lists (e.g., Bink, Marsh, & Hicks, 1999; Hoffman, 1997; Johnson, Raye, Foley, & Foley, 1981). An ANOVA that included list context (medium-with-shallow vs. medium-with-deep), task order (shallower vs. deeper LOP task first), and list attribution (list 1 vs. list 2) revealed a robust interaction of task order with list attribution (i.e., the bias), $F(1,44) = 23.87$, $MSE = .002$, $\eta^2 = .35$, as well as a three-way interaction with group, $F(1,44) = 13.08$, $MSE = .002$, $\eta^2 = .23$. When list 1 was the shallower list, subjects in the medium-with-shallow group were more likely to attribute the source of a new word to list 1 than to list 2 (.09 vs. .03), whereas when list 2 was the shallower list, they were more likely to attribute the source of a new word to list 2 than to list 1 (.12 vs. .02), $F(1,22) = 20.10$, $MSE = .004$, $\eta^2 = .52$ for the interaction. The same pattern was shown by the medium-with-deep group, but the effect was much smaller (.03 vs. .01; .03 vs. .02) and the interaction was only marginally significant, $F(1,22) = 3.98$, $MSE = .0004$, $p = .06$, $\eta^2 = .15$. Finally, unsure judgments to new words did not differ as a function of list context, $F(1,44) = 1.63$, $MSE = .012$, $p = .21$, $\eta^2 = .04$.

Experiment 4

Experiment 1 revealed that list context can influence subjects' remember/know reports, and Experiment 2 suggested that these effects occur only when sets of items that systematically differ in memorability are included on the test list. In Experiment 3, we showed that subjects in both list-context groups appear to be equally able to recollect and utilize information that specifies list source. Following Gruppuso et al. (1997), we take this pattern of results as evidence that the test-list context modulates subjects' functional definitions of remembering and knowing, rather than the contents of their recollections.

[2] The rate of false source attributions was higher than the rate of recognition judgments for new words in Experiments 1–2. We are not sure what to make of this result, given that careful responding was emphasized to the same extent in all of our experiments.

Building on Jacoby, Kelley, and Dywan's (1989) attributional account of memory, Conway and Dewhurst (1995a) claimed that different encoding conditions lead to the retrieval of different attributes at test, which in turn lead subjects to make different attributions about their state of awareness. Here we take this claim one step further by suggesting that recollections for items from the *same* encoding condition (e.g., the medium LOP task) can also be attributed to different states of awareness depending on the test-list context (cf. Whittlesea, 2002b). In Experiment 4, we sought a stronger test of our claim that subjects' definitions of what constitutes remembering are qualitatively different in the medium-with-shallow and medium-with-deep groups.

How might the medium-with-shallow and the medium-with-deep groups' definitions of remembering differ? Consider an example in which the word "clown" was studied on each group's medium LOP list. At test, the medium-with-shallow group might base their remember judgments primarily on recollection of spontaneous thoughts or associations generated during study rather than on recollection of the encoding task with which an item had been studied. This would be an adaptive approach because recollecting having thought "I sure hate clowns" would provide much more distinctive information than recalling "I said clown was a common/uncommon word" given that many words had been judged as common/uncommon. The medium-with-deep group, in contrast, might emphasize recollection of thoughts relating to the deep LOP task as a basis for remember responses. This strategy would be adaptive because subjects would be quite likely to recall their thoughts about whether or not they wanted a particular item on their desert island; such thoughts would be highly distinctive. Failing to recall such memories for "clown," the medium-with-deep group might make a know judgment rather than a remember judgment-even if they, like the medium-with-shallow group, had also thought "I sure hate clowns" during the study phase-because of their emphasis at test on recollecting the encoding task rather than spontaneous thoughts or associations.

The primary goal of Experiment 4 was to examine if subjects' reported strongest recollections would differ as a function of the test-list context. To this end, we repeated the same study phase used in Experiments 1B to 3. At test, subjects were again requested to provide remember/know judgments for recognized items. Following remember judgments subjects made two additional reports. Subjects first rated the extensiveness of their recollection for the current item, to get subjects to actively evaluate the contents of their recollections prior to making their second report. Critically, subjects then reported to the experimenter their strongest recollection about the item. The functional definition account suggests that some kinds of information will be experienced as more distinctive or diagnostic in one list context than in another, and hence that the two list-context groups will report qualitatively different types of recollected attributes as their strongest recollection. As foreshadowed above, the medium-with-shallow group might emphasize recollection of thoughts and associations more than the medium-with-deep group, whereas the medium-with-deep group might emphasize recollection of the encoding task used to study a particular item more than the medium-with-shallow group.

An alternative explanation for the list context effects in Experiment 1 is that the medium-with-shallow group might simply have adopted a more liberal remember/know criterion along a memory strength continuum relative to the medium-with-deep group, because the former group would have had weaker memory traces available to them on average. Were that the case, the medium-with-shallow group might have placed high confidence experiences for which no specific details could be recollected into the remember category, while the medium-with-deep group might have placed equivalent experiences into the know category. Conceptualized in this way, a criterion-setting account would be consistent with the list-context effects observed on remembering and knowing in Experiment 1.

A secondary goal of Experiment 4 was therefore to ensure that the medium-with-shallow group were not simply placing high confidence responses that were not associated with any recollected details into the remember category. We reasoned that requiring subjects to provide the basis they were using for each remember response would ensure that both list-context groups reported remembering only when they recollected details about their encounter with a studied item. If the list-context effects in Experiment 1 were due to the medium-with-shallow groups in those experiments placing strong "know experiences" into the remember category, then list-context effects should be eliminated here. In contrast, on the functional definition account, subjects' definitions of remembering and knowing are constructed based on their recollection experiences during the test list, and hence this account predicts that the list-context effects found in Experiment 1 should be replicated here.

Method

Subjects. Subjects were randomly assigned to the two groups ($n = 24$ each). We replaced one subject who claimed to remember more than 10% of the new words.

Materials, design, and procedure. The materials, design, and study phase procedure were the same as in Experiment 1B. The test phase procedure was also the same as Experiment 1B except that subjects performed two additional tasks after they made a remember judgment.

Table 1
Categories used to classify the "strongest recollection" accompanying remember judgments in Experiment 4

Strongest recollection	Description of recollection
List source	The encoding task or the response that was given
Thought/association	A thought or association cued by the word
Visual image	A visual image constructed for the word
Appearance	Seeing the word on the computer screen
List position	The position of the word within the study list
Emotion	An emotional reaction elicited by the word
No basis given	Failure to provide a recollection

To ensure that subjects evaluated the contents of their recollection prior to reporting their strongest recollection for an item, they first rated "How much recollection" they experienced for the item using the following 4-point scale: 1 = none, 2 = very little, 3 = some, 4 = a lot. These ratings are not discussed further. The following query then appeared on their screen: "What is your strongest recollection about studying this word?" Subjects verbally reported their strongest recollection, which was then classified on-line by the experimenter into one of the seven categories listed in Table 1. An eighth category, "other," was available but was not used.

Results and discussion

Remember/know data

The proportion of remember and know judgments for each set of items for each group in Experiment 4 are presented in Fig. 1. Although subjects were required to provide a basis for each remember judgment here, the pattern of data was very similar to that found in Experiment 1.

Appropriate use of judgments. In the medium-with-shallow and medium-with-deep groups, the respective low rates of remember (.01 vs. .02) and know (.15 vs. .13) judgments to new words were not reliably different, $F\text{s} < 1$.

Level of processing effects. The LOP manipulations at encoding produced the expected effects on overall recognition and on remembering (shallow < medium, medium < deep; all $p\text{s} < .0001$). The effect of LOP on know judgments closely replicated the pattern in Experiment 1: knowing was equally likely for shallow and medium items in the medium-with-shallow group (.29 vs. .29), $F < 1$, but was more likely for medium than for deep items in the medium-with-deep group (.40 vs. .30), $F(1,22) = 13.04$, $MSE = .01$, $\eta^2 = .37$. Thus, whether LOP/duration affected know judgments depended on the pair of groups being compared. Finally, although estimates of familiarity (excluding 0s and 1s) were always greater for the deeper LOP condition within each group, this pattern was reliable only for the medium-with-shallow group (.67 vs. .32), $F(1,17) = 47.85$, $MSE = .02$, $\eta^2 = .74$, and not for the medium-with-deep group (.77 vs. .71), $F(1,14) = 1.28$, $MSE = .02$, $p = .28$, $\eta^2 = .08$.[3]

List-context effects. The small advantage in overall recognition of medium items in the medium-with-shallow group over the medium-with-deep group was nearly the same as that found in Experiment 1, but here, with smaller sample sizes, it did not approach significance (.86 vs. .80), $F(1,44) = 2.34$, $MSE = .02$, $p = .13$, $\eta^2 = .05$. Most importantly, even though subjects were required to provide a basis for each of their remember judgments, the medium-with-shallow group was again much more likely than the medium-with-deep group to classify their recognition of medium items as an experience of remembering (.57 vs. .40), $F(1,44) = 7.33$, $MSE = .05$, $\eta^2 = .14$, and were correspondingly less likely to classify it as an experience of knowing (.29 vs. .40), $F(1,44) = 4.67$, $MSE = .04$, $\eta^2 = .10$. An analysis of the remember:recognize ratio (see Experiment 1) showed that the proportion of recognized medium items that were classified as remember was greater in the medium-with-shallow group than in the medium-with-deep group (.66 vs. .48), $F(1,44) = 7.10$, $MSE = .06$, $\eta^2 = .14$. These findings closely replicate the list-context effects in Experiment 1.

The persistence of list-context effects when subjects were required to provide a basis for each of their remember judgments rules against the possibility that the medium-with-shallow groups in Experiments 1 and 4 classified strong know experiences as experiences of remembering. Requiring subjects to provide a basis for each remember judgment does not appear to have made them more conservative, suggesting that subjects in both experiments used the remember and know categories

[3] Familiarity estimates at floor or ceiling (i.e., 0 or 1) were excluded from analysis. In the medium-with-shallow group, three familiarity estimates for medium items were at ceiling and one was at floor, and two estimates for shallow items were at floor. In the medium-with-deep group, eight familiarity estimates for deep items were at ceiling, and one estimate for medium items was at ceiling. The loss of eight subjects with ceiling familiarity estimates for deep LOP items may explain why the effect of LOP on familiarity estimates was not reliable for the medium-with-deep group.

very carefully. Hence, according to the criterion model, subjects in both experiments must have set their remember/know criterion in a relatively conservative manner. To accommodate our data, proponents of a criterion account could assume that a given piece of recollected information for a medium LOP item has more "strength" in one list context than it does in another. Framed this way, the criterion account is quite compatible with the functional account, according to which list-context effects arise because subjects in the two groups construct different definitions of remembering and knowing based on their experiences during the test list.

Estimates of familiarity were equivalent in the medium-with-shallow and medium-with-deep groups (.64 vs. .68), $F < 1$.[3] This result is consistent with the familiarity process in dual-process theory, under the assumption that both groups were equally conservative (Yonelinas, 2002).

Basis for remember judgments

Subjects reported their strongest recollection for each item classified as remembered. For each type of item within each group, we computed the proportion of remember judgments that were associated with each category of recollection listed in Table 1 (see Fig. 3). The proportions of each type of recollection for a given set of items within each group therefore add up to 1. Several subjects never made a remember judgment to any shallow and/or new items. As a result, the proportions for shallow words in the medium-with-shallow group are based on the 15 subjects who made at least one remember judgment to a shallow word, and the proportions for new words are based on the nine subjects in the medium-with-shallow group and the eight subjects in the medium-with-deep group who made at least one remember judgment to a new word. Thus Fig. 3 displays the *relative* proportion of each basis but note that the *absolute* rate of each basis in a given condition depended on the absolute rate of remembering in that condition.

For brevity, we focus on the main question of interest, namely whether the medium-with-shallow and medium-with-deep groups, who differed in the amount of remember judgments they made to medium items, also differed in the relative frequency with which they reported different types of recollected details about their study experiences. A $7 \times 2 \times 2$ ANOVA with recollection basis, list-context group, and task order as the factors revealed only two significant effects: a main effect of recollection basis, $F(6, 264) = 31.87$, $MSE = .04$, $\eta^2 = .42$, and an interaction of recollection basis with list context, $F(6, 264) = 4.89$, $MSE = .04$, $\eta^2 = .10$.

Two differences emerged in follow-up tests in which we compared the rates with which the two groups reported each recollection basis. The first difference was the relative frequency with which the two groups' strongest recollection involved the list source (e.g., which encoding task the item was studied in, and/or the response that was given for the item in the task at hand). List source recollections were twice as likely to be offered by the medium-with-deep group as by the medium-with-shallow group (.51 vs. .26), $F(1, 44) = 7.32$, $MSE = .10$, $\eta^2 = .14$. The opposite pattern was observed for recollections of thoughts and associations experienced during encoding: the medium-with-shallow group reported thoughts/associations as the basis for their remember judgments for medium items relatively more often than did the medium-with-deep group (.35 vs. .21), $F(1, 44) = 6.54$, $MSE = .04$, $\eta^2 = .13$.

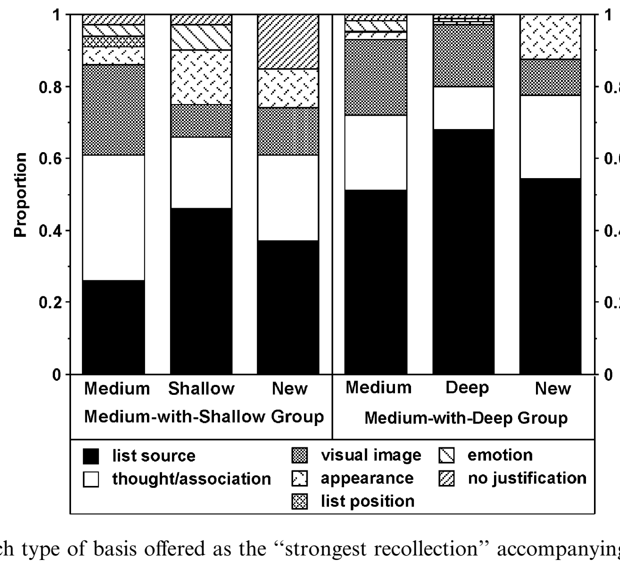

Fig. 3. Mean proportion of each type of basis offered as the "strongest recollection" accompanying remember judgments to studied words and new words in the medium-with-shallow (M/S) and medium-with-deep (M/D) groups in Experiment 4.

There were no reliable differences between the two groups on the other types of recollections.

A second $7 \times 2 \times 2$ ANOVA was performed on the *absolute* proportions of each recollection basis. This analysis revealed main effects of recollection basis, $F(6, 264) = 24.84$, $MSE = .01$, $\eta^2 = .36$, and list context, $F(1, 44) = 7.10$, $MSE = .01$, $\eta^2 = .14$, and more importantly, an interaction of recollection basis with list context, $F(6, 264) = 2.52$, $MSE = .01$, $\eta^2 = .05$. As in the relative-proportion analysis, follow up tests showed that the medium-with-shallow group was more likely than the medium-with-deep group to report thoughts and associations as their strongest recollection (.19 vs. .09), $F(1, 44) = 11.23$, $MSE = .01$, $\eta^2 = .20$. However, there was no difference in the rate with which the two groups reported list source as their strongest recollection (.15 vs. .19), $F < 1$. This latter null effect is consistent with null difference between the groups in their rate of correct source judgments in Experiment 3 and suggests that the medium-with-shallow group's greater rate of remember judgments for medium items relative to the medium-with-deep group may be due more to this group's orientation toward recollecting their thoughts and associations rather than to their inability to recollect that these items were on the medium LOP list. There were no reliable differences between the two groups on the remaining recollection bases.

The medium-with-deep group appears to have emphasized recollection of the list in which an item was studied, a reasonable approach given the unique and memorable nature of the deep LOP task (i.e., the desert island task). As a result, the medium-with-deep group was proportionally more likely than the medium-with-shallow group to report that their strongest recollection for medium LOP items was that these items had appeared on the medium-LOP list (i.e., the common-or-not task), even though the two groups were equally able to recollect information that allowed them to specify list source when they were asked to focus on these attributes (Experiment 3). When the absolute rates were considered, no difference in the proportion of list-source reports made for remembered medium items was found between the two groups (keeping in mind that the rate of remembering was lower in the medium-with-deep group). In contrast, the medium-with-shallow group appears to have emphasized recollection of thoughts or associations formed in connection with items more than the medium-with-deep group, whether measured in relative or absolute terms. This too would be an adaptive approach, because recollecting a thought or association about an item would be more distinctive than recollecting whether a common/uncommon judgment had been made about an item at study. Test-list context therefore appears to influence the relative emphasis subjects place on qualitatively different types of details when making their remember/know judgments.

General discussion

Whether subjects report remembering or knowing that they studied members of a given set of items depends, in part, on how they evaluate their recognition experiences for a second set of studied items on the test list. We found that medium LOP items were more likely to be classified as remembered, and less likely to be classified as known, when mixed with shallow rather than with deep LOP items, even though overall recognition was not reliably affected by list context (Experiments 1 and 4). However, a pooling of the data from Experiments 1A, B, and 4 did reveal a modest overall recognition advantage for the medium-with-shallow group, $F(1, 140) = 6.13$, $MSE = .02$, $\eta^2 = .04$, which did not interact with task order, $F < 1$.[4] Although manipulating list context using an LOP manipulation thus appears to have a small effect on *whether* items are recognized (unlike the null effects reported with duration-based or repetition-based list strength manipulations; Hirshman, 1995; Murnane & Shiffrin, 1991a, 1991b; Ratcliff et al., 1990, 1992; Yonelinas et al., 1992), it clearly has a much more robust effect on *how* items are recognized. Remember/know judgments can therefore reveal large effects of manipulations that would often remain undetected if only yes/no recognition judgments were used (e.g., Gardiner, Gregg, Mashru, & Thaman, 2001; Gardiner & Richardson-Klavehn, 2000; Macken, 2002).

List-context effects disappeared when, after studying items in two LOP tasks, only items from a single LOP list were encountered on the recognition test (Experiment 2) or when the task required subjects to focus on those aspects of their recollection that enabled them to specify the list in which an item had been studied (Experiment 3). These results are consistent with our claim that list-context effects are not mediated by differences in the amount of list-related information that is recollected for a given set of items. Rather, we suggest that when subjects experience items from different encoding situations on the test list, and are not asked to focus on a particular subset of the attributes that they recollect about these items, they will come to use different attributes to define remembering and knowing for themselves (e.g., Dewhurst & Conway, 1994; Gruppuso et al., 1997; Whittlesea, 2002a, 2002b). These differences in functional definitions arise even when the remember/

[4] Although there was a small advantage in the overall hit rate for medium items in the medium-with-shallow group, this group also had a higher overall false alarm rate, relative to the medium-with-deep group. Measures of discriminability for medium items, based on overall recognition (remember + know), were the same in the medium-with-shallow and medium-with-deep groups, whether computed using A (.91 vs. .91), $F < 1$, or hit rate minus false alarm rate (.70 vs. .71), $F < 1$, and neither interacted with task order, $F\text{s} < 1$.

know instructions provide examples of several types of attributes that qualify as experiences of remembering (e.g., thoughts, images, and feelings). Experiment 4 provided strong support for the functional account by showing that the medium-with-shallow group placed relatively more emphasis on recollected thoughts/associations (presumably because these details would feel more distinctive than list source information), whereas the medium-with-deep group placed relatively more emphasis on recollections that specified list membership (presumably because these details would be distinctive for items that had appeared on the desert-island task). We elaborate on this functional account below, after we review the implications of our results for a variety of other accounts of the distinction between remembering and knowing.

Implications of list-context effects for accounts of remembering and knowing

Remembering and knowing as indexes of stable states of awareness or of separate memory systems. The influence of test-list context on remembering and knowing cannot readily be accommodated by an account that attributes remember responses to activation of an episodic memory system and know responses to activation of a semantic or perceptual memory system (e.g., Gardiner & Java, 1993). Our findings are not consistent with the idea that the effects of list context on remembering were due to differential encoding or retrieval from episodic memory (e.g., compare Experiments 1 vs. 2). Moreover, it is not clear why impaired episodic remembering of medium LOP items in the medium-with-deep group would facilitate the operation of a semantic or perceptual memory system and thereby increase know responses, yet in Experiments 1 and 4 we found that reports of knowing for medium LOP items increased in list contexts in which reports of remembering decreased. Bowler, Gardiner, and Grice (2000) suggested that under some circumstances subjects might alter their degree of reliance on semantic versus episodic memory systems. Applied here, one might argue that use of a semantic or perceptual memory system increases only when use of the episodic memory system fails. However, it seems counterintuitive to claim that subjects in the medium-with-deep group, who should enjoy the richest set of experiences of remembering, would rely on their episodic memory systems less than subjects in the medium-with-shallow group. Our results, as well as those of others (e.g., Dewhurst & Conway, 1994; Dewhurst & Parry, 2000; Gruppuso et al., 1997; Whittlesea, 2002b; Xu & Bellezza, 2001), suggest that remember/know judgments emerge from an attributional process rather than serving as reliable and direct indexes of stable response categories (Gardiner & Richardson-Klavehn, 2000) or memory systems (Tulving, 1985).

Remembering and knowing as indicators of processing distinctiveness/salience and fluency. Rajaram (1996, 1998) has developed an account in which remembering is enhanced by cognitive operations that increase the distinctiveness or salience of processing, whereas knowing is enhanced by operations that increase processing fluency. Earlier work suggested a link between remembering and conceptual processing, and between knowing and perceptual processing (e.g., Rajaram, 1993; Rajaram & Roediger, 1997), but more recent work has shown that processing distinctiveness and processing fluency can each be increased through either perceptual or conceptual means, which can be manipulated either at encoding or at test (e.g., Rajaram, 1996, 1998; Rajaram & Geraci, 2000). The test-list context effect on remember judgments to medium items in Experiments 1 and 4 can be accommodated by this account if it is assumed that medium items were processed more distinctively or saliently when tested in the presence of shallow items than when tested in the presence of deep items. The reverse effect of list context on know judgments in these experiments would appear to imply that medium items were processed less fluently when tested in the presence of shallow items than when tested in the presence of deep items (e.g., perhaps because the recognition test was generally more difficult in the former case). Framed in this way, the current results are compatible with the distinctiveness/fluency account, although it would be useful for future research to provide additional support for the claims that medium items are processed more distinctively/saliently, yet less fluently, in the shallow-item context than in the deep-item context.

Remembering and knowing as correlates of dual-process mechanisms. Proponents of a dual-process account (e.g., Jacoby, 1991; Yonelinas, 2002) could propose that a list context consisting of more memorable items (medium-with-deep group) rather than less memorable items (medium-with-shallow group) does not affect the familiarity process. The absence of list context effects on estimates of familiarity reported in Experiments 1 and 4 is consistent with this proposal. The effects of list context on remember judgments, which are thought to map onto a threshold recollection process, may be more difficult for dual-process models to accommodate. For example, it might be argued that certain list contexts can interfere with the recollection process. One challenge for this explanation is the failure of prior research (e.g., Hirshman, 1995; Murnane & Shiffrin, 1991a, 1991b; Ratcliff et al., 1990, 1992; Yonelinas et al., 1992) and our research (Experiments 1–4) to obtain reliable list-strength effects on overall recognition memory performance. Although a pooling of the data from Experiments 1 and 4 revealed a significant list-strength effect on overall recognition, it is unclear whether such a small and non-robust effect on overall recognition can explain the large and robust effects of list context on remembering and knowing in these experiments. In addition, the dual-process account will need to explain how the weighting of different recollected attributes is influenced by list context (Experiment 4).

Remembering and knowing as different signal detection criteria. It has been suggested that manipulations that affect remember/know judgments without affecting overall recognition provide strong evidence against a single-process dual-criterion criterion model (e.g., Bowler et al., 2000; Conway & Dewhurst, 1995b; Conway, Dewhurst, Pearson, & Sapute, 2001; Gardiner et al., 2001). However, the criterion model can mimic our list-context effects if it is assumed that both list-context groups set a very similar yes/no criterion, but the medium-with-deep group set a more conservative remember/know criterion (e.g., Hirshman & Master, 1997). Unfortunately, this approach does not provide any insight into how and why the criteria would come to be set in these particular ways in each group. For example, why would the medium-with-deep group, who place a more conservative remember/know criterion, not also place a more conservative yes/no criterion? Thus, a criterion-based account, even if it can mimic many patterns of data, does little to help shape our understanding of the experiential states of awareness that accompany recognition (Whittlesea, 2002a). Moreover, criterion accounts will have to be elaborated to explain why list context systematically influenced the types of qualitative details subjects tended to report as their strongest recollection for medium items in Experiment 4.[5]

Remembering and knowing as context-dependent and functionally defined experiences. We think that the framework developed by Gruppuso et al. (1997) in their studies of estimates of recollection and familiarity obtained with Jacoby's (1991) process-dissociation procedure can be fruitfully applied to our results as well. Gruppuso et al. proposed that the psychological constructs of recollection and familiarity should be defined in functionalist terms. Recollection occurs when information recruited from memory enables the rememberer to perform the task at hand (e.g., recalling in which of two lists a word had been studied, or identifying a familiar person encountered in an unusual context). Familiarity occurs when information recruited from memory makes an item seem familiar, but does not allow the subject to sufficiently specify the source of their memory given the task at hand. Importantly, the claim made here (and by Dewhurst & Conway, 1994) is not that people merely *respond differently* on remember/know or inclusion/exclusion tasks as a consequence of whether or not they recollect the sorts of information that enable them to perform the discrimination at hand, but that they also have *different phenomenological experiences* (i.e., nagging familiarity vs. satisfying recollection). From this perspective, some memory information may be recollected about the butcher on the bus—enough to produce an agonizing feeling of familiarity—but this information is too generic, incomplete, or vague to enable a unique identification of the individual (see Koriat, 1995, for an analogous account of tip-of-the-tongue states and feeling-of-knowing judgments).

The functional account combines aspects of the experiential states, transfer-appropriate processing, and dual-process approaches. Different experiential states arise depending on whether the information recruited from memory is sufficiently detailed and source specifying to warrant an attribution to a particular source in one's past experience (e.g., Conway & Dewhurst, 1995a; Whittlesea, 2002b), which in turn depends on specifics of the study/test context (e.g., Morris, Bransford, & Franks, 1977; Tulving & Thomson, 1973). Moreover, consistent with both separate-systems and dual-process views, memory information that typically contributes to recollection tends to be relatively conceptual, idiosyncratic, and integrative, whereas memory information that typically contributes to familiarity tends to be relatively perceptual/semantic and generic. But, again, the kinds of memory information that support recollection versus familiarity depend on the study (e.g., LOP) and test (e.g., list context) conditions (Experiment 4). This is not because each state arises from a unique memory system or process, but rather because some types of recollected information are more distinctive and source specifying than others in a given context.

This account fits well with Perfect, Mayes, Downes, and Van Eijk's (1996, Experiment 5) finding that 6% of remember responses were not associated with recollection of contextual information, and, more surprisingly, that 32% of know responses *were* accompanied by recollection of some form of contextual knowledge.

[5] Another purported challenge to the criterion model may be the common finding that measures of discriminability (e.g., A') are higher when computed based on overall recognition (remember + know) scores than when computed from remember scores alone (e.g., Gardiner et al., 2001). This pattern has been taken as evidence that know judgments provide information incremental to that provided by remember judgments (e.g., Gardiner & Conway, 1999). For the record, in Experiments 1 and 4, we also found that estimates of A were greater when computed based on overall recognition scores rather than remember scores alone: for shallow items (.76 vs. .74), $F(1,83) = 5.71$, $MSE = .002$, $\eta^2 = .06$, medium items (.92 vs. .87), $F(1,94) = 71.26$, $MSE = .002$, $\eta^2 = .43$, and deep items (.94 vs. .91), $F(1,94) = 41.10$, $MSE = .001$, $\eta^2 = .30$ in Experiment 1, and for medium items (.90 vs. .86), $F(1,47) = 22.60$, $MSE = .002$, $\eta^2 = .32$, and deep items (.94 vs. .91), $F(1,23) = 9.79$, $MSE = .002$, $\eta^2 = .30$ in Experiment 4 ($F < 1$ for shallow items, but $n = 15$). (Note that some subjects never made a remember judgment for shallow items.) Hirshman (1998) suggested that small differences in A' such as these may simply reflect measurement error.

Regarding the latter result, the authors wrote "this seems to us to represent evidence of a continuum of contextual knowledge along which subjects are required to draw a distinction between recollection and knowing" (p. 810). We agree with Perfect et al., although rather than viewing the recollected attributes as a single continuum with a more or less arbitrary criterion, we suggest that different attributes will be used to support remember or know judgments depending on their diagnosticity in the current context (cf. Experiment 3 vs. 4).

We have suggested that list context influences how subjects define remembering. Note that differences in subjects' definitions of what constitutes a sufficient basis for remembering might influence not only how subjects experience memory information that comes to mind in response to a test probe, but also the specific types of memory information that come to mind (e.g., subjects may set up different self-generated retrieval cues in different test-list contexts). Both mechanisms are compatible with the functionalist account, but a goal for future research is to assess their relative contributions to the effect of test-list context on judgments of remembering and knowing.

We suggest that the functionalist account may provide a psychological framework for the criterion model account of the remember/know distinction. Estimates of bias, sensitivity, and response criteria obtained with signal-detection theory analyses are not psychological mechanisms, rather, they are summary-level descriptions of data (e.g., Whittlesea, 2002a). Where a particular piece of memory information falls on the weak–strong continuum of the signal-detection theory metaphor depends on the diagnostic utility of that piece of information as evidence of "oldness," which in turn depends on the specifics of the study/test situation. That is, memory strength, as described in signal-detection theory, is not a primitive property of the memory response to an individual item in isolation, but rather is an emergent property of the quantitative and qualitative content of that memory response in the present context (Donaldson et al., 1996; Xu & Bellezza, 2001). Thus a memory response that gives rise to a subjective experience of recollection is not simply "more of the same stuff" that gives rise to an experience of familiarity or knowing, as subjects' recollection reports in Experiment 4 show. Rather, a memory response is experienced as remembering if, in the study/test context, it is experienced as sufficiently and uniquely specifying a moment in one's past experience. From this perspective, a dual-criterion signal-detection model is merely a way of describing the data that emerge when people make a relatively gross old/new discrimination and then further decide whether or not their recognition gives rise to an experience of remembering a unique moment from their personal history.

Relevance beyond word lists: A final note

We believe our results also have an important practical implication, namely that in any multi-question interview situation, the likelihood that people will claim to remember something from their past may depend, in part, on how well they remember the answers to the other questions posed in the interview. For example, imagine a reminiscer is looking through a photo album laden either with photos from favorite family events (e.g., a parent's photo album), or with photos from more obscure family events (e.g., a distant relative's photo album). Our data suggest that the reminiscer may claim to remember a particular event which stands out against a background of less memorable photos, but may only claim to know this event from his/her past if it is presented in an album of family favorites. The ramifications of list-context effects on recognition judgments become rather sobering if the photo album from this example were instead a series of photographs from a crime scene, and the reminiscer a key eyewitness. The list-context paradigm used here can easily be adapted for studying recognition memory in situations such as these.

Acknowledgments

This research was supported by the Natural Sciences and Engineering Research Council of Canada through research grants to Glen Bodner and Steve Lindsay. We thank Shannon Skov-Rackette for his unstinting assistance with data collection, and Suparna Rajaram, Phil Higham, Andy Yonelinas, and an anonymous reviewer for providing helpful comments on earlier versions of the manuscript.

References

Bink, M. L., Marsh, R. L., & Hicks, J. L. (1999). An alternative conceptualization to memory "strength" in reality monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 804–809.

Bowler, D. M., Gardiner, J. M., & Grice, S. J. (2000). Episodic memory and remembering in adults with Asperger Syndrome. *Journal of Autism and Developmental Disorders, 30*, 295–304.

Conway, M. A., & Dewhurst, S. A. (1995a). Remembering, familiarity, and source monitoring. *Quarterly Journal of Experimental Psychology, 48*, 125–140.

Conway, M. A., & Dewhurst, S. A. (1995b). The self and recollective experience. *Applied Cognitive Psychology, 9*, 1–19.

Conway, M. A., Dewhurst, S. A., Pearson, N., & Sapute, A. (2001). The self and recollection reconsidered: How a 'failure to replicate' failed and why trace strength accounts of recollection are untenable. *Applied Cognitive Psychology, 15*, 1–14.

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

Dewhurst, S. A., & Conway, M. A. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1088–1098.

Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness, and recollective experience. *European Journal of Cognitive Psychology, 12*, 541–551.

Dodson, C. S., & Johnson, M. K. (1996). Some problems with the process-dissociation approach to memory. *Journal of Experimental Psychology: General, 125*, 181–194.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition, 24*, 523–533.

Donaldson, W., MacKenzie, T. M., & Underhill, C. F. (1996). A comparison of recollective memory and source monitoring. *Psychonomic Bulletin & Review, 3*, 486–490.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition, 16*, 309–313.

Gardiner, J. M., & Conway, M. A. (1999). Levels of awareness and varieties of experience. In B. H. Challis & B. M. Velichovsky (Eds.), *Stratification in Cognition and Consciousness* (pp. 237–254). Amsterdam: John Benjamin.

Gardiner, J. M., & Gregg, V. H. (1997). Recognition memory with little or no remembering: Implications for a detection model. *Psychonomic Bulletin & Review, 4*, 474–479.

Gardiner, J. M., Gregg, V. H., Mashru, R., & Thaman, M. (2001). Impact of encoding depth on awareness of perceptual effects in recognition memory. *Memory & Cognition, 29*, 433–440.

Gardiner, J. M., & Java, R. I. (1993). Recognition memory and awareness: An experiential approach. *European Journal of Cognitive Psychology, 5*, 337–346.

Gardiner, J. M., Java, R. I., & Richardson-Klavehn, A. (1996). How levels of processing really influences awareness in recognition memory. *Canadian Journal of Experimental Psychology, 50*, 114–122.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. I. M. Craik (Eds.), *Handbook of memory* (pp. 229–244). New York: Oxford University Press.

Gardiner, J. M., Richardson-Klavehn, A., & Ramponi, C. (1998). Limitations of the signal detection model of the remember-know paradigm: A reply to Hirshman. *Consciousness and Cognition, 7*, 285–288.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1–67.

Gruppuso, V., Lindsay, D. S., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 259–278.

Higham, P. A., & Vokey, J. R. (2002). The implications of illusory recollective experiences for dual-process and fuzzy trace models of recognition memory. Submitted.

Hintzman, D. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*, 411–428.

Hirshman, E. (1995). Decision processes in recognition memory: Criterion shifts and the list-strength paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 302–313.

Hirshman, E. (1998). On the utility of the signal detection model of the remember-know paradigm. *Consciousness and Cognition, 7*, 103–107.

Hirshman, E., & Henzler, A. (1998). The role of decision processes in conscious recollection. *Psychological Science, 9*, 61–65.

Hirshman, E., & Master, S. (1997). Modeling the conscious correlates of recognition memory: Reflections on the remember/know paradigm. *Memory & Cognition, 25*, 345–351.

Hoffman, H. G. (1997). Role of memory strength in reality monitoring decisions: Evidence from source attribution biases. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 371–383.

Inoue, C., & Bellezza, F. S. (1998). The detection model of recognition using know and remember judgments. *Memory & Cognition, 26*, 299–308.

Jacoby, L. L. (1991). A process-dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.

Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 3–26.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*, 306–340.

Jacoby, L. L., & Kelley, C. M. (1991). Unconscious influences of memory: Dissociations and automaticity. In D. Milner & M. Rugg (Eds.), *The neuropsychology of consciousness* (pp. 201–233). London: Academic Press.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–421). Hillside, NJ: Lawrence Erlbaum Associates.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1997). The relationship between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to the study of consciousness* (pp. 13–47). Mahwah, NJ: Erlbaum.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3–28.

Johnson, M. K., Raye, C. L., Foley, H. J., & Foley, M. A. (1981). Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology, 94*, 37–64.

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124*, 311–333.

Kučera, J., & Francis, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.

Lindsay, D. S., & Johnson, M. K. (1991). Recognition memory and source monitoring. *Bulletin of the Psychonomic Society, 29*, 203–205.

Macken, W. J. (2002). Environmental context and recognition: The role of recollection and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 153–161.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*, 252–271.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer-appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519–533.

Mulligan, N. W., & Hirshman, E. (1997). Measuring the bases of recognition memory: An investigation of the process-dissociation framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 280–304.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89*, 609–626.

Murnane, K., & Shiffrin, R. M. (1991a). Interference and the representation of events in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 855–874.

Murnane, K., & Shiffrin, R. M. (1991b). Word repetitions in sentence recognition. *Memory & Cognition, 19*, 119–130.

Perfect, T. J., Mayes, A. R., Downes, J. J., & Van Eijk, R. (1996). Does context discriminate recollection from familiarity in recognition memory? *Quarterly Journal of Experimental Psychology, 49A*, 797–813.

Rajaram, S. (1993). Remembering and knowing: Two ways of accessing the personal past. *Memory & Cognition, 21*, 89–102.

Rajaram, S. (1996). Perceptual effects on remembering: Recollective processes in picture recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 365–377.

Rajaram, S. (1998). Conceptual and perceptual effects on remembering: The role of salience/distinctiveness. *Psychonomic Bulletin & Review, 5*, 71–78.

Rajaram, S., & Geraci, L. (2000). Conceptual fluency selectively influences knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1070–1074.

Rajaram, S., & Roediger, H. M. (1997). Remembering and knowing as states of consciousness during retrieval. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to the study of consciousness* (pp. 213–240). Mahwah, NJ: Erlbaum.

Ratcliff, R., Clark, S., & Shiffrin, R. M. (1990). The list-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 163–178.

Ratcliff, R., Sheu, C., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review, 99*, 518–535.

Rugg, M. D., Schloerscheidt, A. M., & Mark, R. E. (1998). An electrophysiological comparison of two indices of recollection. *Journal of Memory and Language, 39*, 47–69.

Shiffrin, R. M., Ratcliff, R., & Clark, S. E. (1990). List-strength effect: II. Theoretical mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 179–195.

Strack, F., & Förster, J. (1995). Reporting recollective experiences: Direct access to memory systems? *Psychological Science, 6*, 352–358.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26*, 1–12.

Tulving, E., & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.

Whittlesea, B. W. A. (2002a). False memory and the discrepancy-attribution hypothesis: The prototype-familiarity illusion. *Journal of Experimental Psychology: General, 131*, 96–115.

Whittlesea, B. W. A. (2002b). Two routes to remembering (and another to remembering NOT). *Journal of Experimental Psychology: General, 131*, 325–348.

Xu, M., & Bellezza, F. S. (2001). A comparison of the multimemory and detection theories of know and remember recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1197–1210.

Yonelinas, A. P. (2001). Components of episodic memory: The contribution of recollection and familiarity. *Philosophical Transcripts of the Royal Society of London, 356*, 1363–1374.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441–517.

Yonelinas, A. P., Hockley, W. E., & Murdock, B. B. (1992). Tests of the list-strength effect in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 345–355.

Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language, 34*, 622–643.

Yonelinas, A. P., & Jacoby, L. L. (1996). Response bias and the process-dissociation procedure. *Journal of Experimental Psychology: General, 125*, 422–434.