Journal of Experimental Psychology:
Learning, Memory, and Cognition
2003, Vol. 29, No. 6, 1095–1105

0278-7393/03/$12.00 DOI: 10.1037/0278-7393.29.6.1095

What Do We Know About What We Cannot Remember? Accessing the Semantic Attributes of Words That Cannot Be Recalled

Asher Koriat, Ravit Levy-Sadot, Eyal Edry, and Sigal de Marcas
University of Haifa

Two experiments examined access to the semantic attributes of words that participants failed to retrieve. The results indicated access to all 3 dimensions of the semantic differential—evaluation, potency, and activity, as revealed by attribute judgments and by the nature of the commission errors made. There was no evidence for superior access to the emotional-evaluative dimension, inconsistent with what may be expected from the claimed primacy of emotion. In comparison with complete recall, partial recall exhibited a slower rate of forgetting and a stronger tendency to elicit know rather than remember responses. The results were discussed in terms of the processes that lead to partial recall and in terms of the possibility that the affective primacy hypothesis does not apply to memory retrieval.

Many observations suggest that when people cannot retrieve a solicited target from memory they can still provide some partial information about it. Consider the following episode concerning the attempt to recall the name of a restaurant:[1]

> *My wife and I were strolling in a neighborhood a bit far from our home. At one stage we recalled that there was a restaurant at the next corner. We had eaten there only once and were unable to recall the name of the restaurant.*
>
> *My wife said to me: I know, it is a name of a composer.*
>
> *I said: I can't remember the name but I think it begins with an M.*
>
> *Her response was: I'm not sure but I think the name is "Beethoven."*
>
> *My immediate response was: That's it, its name is "Mozart."*
>
> *At first I felt quite sure about this, but as we walked along I began feeling less sure, and lo and behold when we finally got to the restaurant we discovered that its name was "Amadeus!"*

This episode illustrates some of the issues addressed in this study. First, it is clear that retrieval is not an all-or-nothing process. Rather, partial information about a solicited memory target can be accessed even when the retrieval of the target fails (R. W. Brown & McNeill, 1966). Second, the partial information that comes to mind may include metacognitive feelings ("At first I felt quite sure . . .") as well as substantive clues about features of the target (". . . it begins with an *M*"). Third, the substantive partial clues fall into two categories: structural-phonological clues (". . . it begins with an *M*") and semantic clues (". . . it is a name of a composer"). Finally, the example illustrates a distinction between two modes of partial access, a direct mode and an indirect mode. In the direct mode, the person deliberately searches for partial clues and is aware of producing only partial information about the elusive target (". . . it begins with an *M*," ". . . it is a name of a composer"). In contrast, in the indirect mode, the kind of partial information stemming from the inaccessible target can only be inferred (by the experimenter) from various aspects of performance, such as the commission errors made. Such errors (e.g., ". . . the name is Beethoven," or ". . . its name is Mozart") reflect one way in which partial information about an elusive target leaks out.

In this study, we examine the type of semantic partial information that is accessible about a word that one fails to recall and explore certain characteristics that are hypothesized to distinguish partial recall from the retrieval of the word itself.

Previous Work on Generic Recall

The best-known research on partial recall is R. W. Brown and McNeill's (1966) classic study of the tip-of-the tongue (TOT). Participants who failed to retrieve a word in response to its definition, but signaled that it was on the TOT, were able to guess some of its structural features. This finding has been replicated and extended in many subsequent studies (see Schwartz, 2002, for a review).

In R. W. Brown and McNeill's (1966) study, as well as in many subsequent TOT studies, the memory pointers (see Koriat & Lieblich, 1977) used to cue the target were word definitions. Because these definitions contain much of the semantic information about the word, it is not surprising that the partial information that participants provided concerned primarily structural-phonological attributes. In contrast, other studies that used different types of pointers demonstrated that participants can also access partial

Asher Koriat, Ravit Levy-Sadot, Eyal Edry, and Sigal de Marcas, Department of Psychology, University of Haifa, Haifa, Israel.

Ravit Levy-Sadot is now at the Department of Psychology, New York University.

A report of Experiment 1 was presented at the XIth Conference of the European Society for Cognitive Psychology, Ghent, September 1999. Support for the preparation of this article by the German Federal Ministry of Education and Research (BMBF) within the framework of German-Israeli Project Cooperation (DIP) is gratefully acknowledged. We thank Limor Sheffer for the analyses of the data.

Correspondence concerning this article should be addressed to Asher Koriat, Department of Psychology, University of Haifa, Haifa 31905, Israel. E-mail: akoriat@research.haifa.ac.il

[1] We are indebted to Joel Norman for relating this episode to us.

information pertaining to semantic and associative aspects of the target (Lovelace, 1987; Yarmey, 1973). Of particular relevance to the present study are the results of Yavuz and Bousfield (1959), which indicated that participants who failed to recall the English translation of a Turkish word were accurate in rating it on the evaluative dimension of the semantic differential (Osgood, 1952). Similar results were obtained by Schacter and Worling (1985) and Koriat (1993).

In contrast to the studies just mentioned, in which memory for partial information was explicitly tested, a study by Nelson, Fehling, and Moore-Glascock (1979) can be seen to concern indirect access to partial information. Participants who learned a list of number–word pairs (e.g., *12–hates*) but failed to retrieve the target exhibited memory saving in learning semantically related pairs. In particular, they demonstrated memory saving for both superordinate and subordinate information. In addition to memory saving, other memory measures that can disclose indirect access to the semantic attributes of words include commission errors (Koriat, 1993; Roediger & McDermott, 1995), false recognition (e.g., Anisfeld & Knapp, 1968), and transfer (Wickens & Cermak, 1967).

In the experiments that follow, we examined the nature of the semantic information that people can access about a momentarily inaccessible word. Participants learned the Hebrew translations of pseudo-Somali words and were tested by having to recall the Hebrew word in response to the Somali cue. When they failed, they were asked to judge its meaning with respect to one of the three dimensions of the semantic differential (Osgood, 1952)—evaluation (good–bad), potency (strong–weak), and activity (active–passive). Explicit access to semantic attributes of the word was inferred from the accuracy of these judgments. In addition, indirect access was inferred from the tendency to make commission errors that have the same polarity on the respective dimension as the correct word. This procedure helped achieve two objectives. The primary objective was to determine which semantic attributes of an elusive memory target are accessible. A secondary objective was to test two characteristics that are expected to distinguish partial recall from full recall.

With regard to the primary objective, an interesting question is whether Zajonc's (1980, 1984) thesis about the primacy of emotion applies also to the retrieval of information from memory. Zajonc argued that people have privileged access to the positive or negative emotional tone of a word and can react affectively to it even before identifying it (see Duckworth, Bargh, Garcia, & Chaiken, 2002). Although this hypothesis was seen by Zajonc to apply to memory retrieval as well, much of the experimental work in support of it comes from studies indicating that participants can pick up the emotional content of briefly presented stimuli that they fail to identify (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Murphy & Zajonc, 1993; Niedenthal, 1990).

Of direct relevance to the present study is work by Bargh, Litt, Pratto, and Spielman (1989). Participants were successful in deciding whether a subliminally presented word had a positive or a negative valence, but not whether it was a synonym of another word. Whereas these results suggest privileged explicit access to the emotional tone of a word, a subsequent study by Bargh, Raymond, and Chaiken (1995; see also Bargh, 1997) suggests that the affective primacy hypothesis also holds true for implicit access: Priming effects were found only between words that had the same polarity on the evaluative dimension of the semantic differential (e.g., when both words had a "good" connotation), but not between words that had similar polarities either on the activity dimension or on the potency dimension.

If similar processes underlie partial access to unaware targets in both perception and memory, then we should expect a similar privileged access to the evaluative dimension in retrieving information from memory. Note that the few studies that examined memory for the semantic attributes of unrecalled words (Koriat, 1993; Schacter & Worling, 1985; Yavuz & Bousfield, 1959) have all focused on the evaluative-emotional attribute of these words, perhaps reflecting an endorsement of the affective primacy hypothesis with regard to memory. In our experiments, we included all three dimensions of the semantic differential. It should be noted that Eysenck (1979), who asked participants to judge rare English words on all three dimensions, found that when participants failed to define the word and also felt that they did not know its meaning, their judgments were accurate only for the evaluative dimension. The question, then, is whether this is also true for judgments made about a memory target that one fails to recall.

The secondary objective of this study was to explore two characteristics of partial recall that follow from the conditions that are assumed to give rise to retrieval blockage. Concurring with fuzzy-trace theory, we assume that forgetting involves not only a loss of some of the features of a memory trace but also a weakening of the bonds that connect the features (Belli, Windschitl, McCarthy, & Winfrey, 1992; Brainerd, Reyna, Howe, & Kingma, 1990). Thus, assuming that information about a word is represented in memory as a collection of attributes (Underwood, 1969), then even when the word becomes inaccessible, information about some of its individual features may survive, giving rise to partial recall. The first prediction that follows from this proposal is that attribute recall should exhibit a slower forgetting rate than item recall. Indeed, several previous studies reported a shallower rate of forgetting for partial than for complete recall, although all of these studies have focused on the fine-grained versus coarse-grained nature of the information that can be retrieved (e.g., Brainerd & Reyna, 1993; Cohen, Stanhope, & Conway, 1992; Dorfman & Mandler, 1994; Kintsch, Welsch, Schmalhofer, & Zimny, 1990).

The second prediction follows from an idea that is prevalent in recent discussions of memory errors: "Inadequate feature binding can result in source memory failure, where people retrieve fragments of an episode but are unable to recollect how or when the fragments were acquired" (Schacter, Norman, & Koutstaal, 1998, p. 291). We propose that partial clues about an elusive memory target (e.g., that the elusive word contains the letter *b* or that it has a negative feeling tone) carry insufficient source information. Therefore, we expect partial recall to be associated more often with know than with remember states of awareness (see Gardiner & Richardson-Klavehn, 2000), whereas item recall may exhibit the opposite pattern. Thus, in Experiment 2 we investigate both the metacognitive experience associated with attribute and item recall as well as the diagnostic value of that experience in distinguishing between correct and wrong responses.

Experiment 1

Experiment 1 focused on the three dimensions of the semantic differential (Osgood, 1952), examining whether the information accessed about an elusive word is confined to the evaluative connotations or is multifaceted. Participants learned the Hebrew translations of pseudo-Somali words. When they failed, at test, to recall the Hebrew word in response to the Somali word, they were asked to judge whether it was good or bad, strong or weak, active or passive. The accuracy of these judgments served as one measure of partial access to semantic information. A second measure was based on the tendency of participants to make commission errors that had the same polarity on the respective dimension as the target word (e.g., responding "happy" instead of "health").

Retention interval was manipulated between participants (10 min vs. 1 week) to test the hypothesis that access to partial information exhibits a slower forgetting rate than the recall of the words proper. In addition, to allow a wider range of variation in full and partial recall, half the participants in each group received three study blocks, and the remaining participants received four study blocks.

Method

Participants. Sixty Hebrew-speaking University of Haifa students (43 women and 17 men) participated in the experiment for course credit.

Stimulus materials. A list of 48 Hebrew words was developed, 8 representing each of the two poles of the three major dimensions of the semantic differential. These words were selected as follows: First, 150 words were selected on the basis of the English norms (Heise, 1965) so that roughly 25 words had extreme ratings on each pole of each of the three dimensions. Ten Hebrew-speaking students then rated the Hebrew translations of these words on six semantic differential scales, two representing each of the three dimensions, good–bad and pleasant–unpleasant for the evaluative dimension, strong–weak and tough–tender for the potency dimension, and active–passive and lively–still for the activity dimension. Forty-eight Hebrew words were then selected on the basis of these ratings (see the Appendix) according to the following criteria: First, eight words represented each of the poles; they had high ratings on that pole (e.g., active), whereas their ratings on the two other dimensions were as close as possible to the point of neutrality. Second, the extremeness of the rating (i.e., the average deviation from the center of the scale on the pertinent dimension) was about the same across the six groups of words representing each of the poles.

In parallel, 48 pseudo-Somali words were invented.[2] These words were one- to three-syllable pronounceable nonsense strings that evoked little definite associations among Hebrew speakers. They were randomly paired with the Hebrew words for each participant.

Apparatus and procedure. The experiment was controlled by a computer. In the first block of the study phase, the Somali–Hebrew pairs were presented for 5 s each, with a 1-s interval between pairs. The Somali word, in Latin letters, appeared on the left side, and the Hebrew word appeared on the right side. In the remaining study blocks, only the Somali word was shown, and participants had to say aloud the Hebrew translation. The corresponding Hebrew word was then presented next to the Somali word (for 2.5 s) either 1 s after the participant responded or after 8 s if no response had been supplied. Presentation order was random for each participant and block.

Thirty participants were assigned to the 10-min (immediate) test condition, and 30 to the 1-week (delayed) test condition. Half the participants in each group were presented with three study blocks, and the other half with four study blocks. For the immediate-test condition, participants were given two filler tasks during the 10-min interval. Participants in the delayed-test group were dismissed and scheduled for the second session 1 week later on the pretext that they would be asked to learn a second language.

The first block of the test phase was similar to the study blocks, except that only the Somali words were presented. When participants failed to supply an answer within the 8-s time limit, they were presented with the attribute-identification task: They were asked to judge the polarity of the nonrecalled word on the pertinent attribute by using a 4-point scale. For the Good–Bad scale, the options were labeled *sure-bad*, *guess-bad*, *guess-good*, and *sure-good*. Similar labels were used for the Weak–Strong and Passive–Active scales. In a second test phase, only items for which participants made a commission error on the first test phase were presented, and participants were only asked to perform the attribute-identification task.

Results

Recall performance. It is important to check the results for recall because several previous studies have demonstrated systematic differences in recall between affective words and neutral words (see LaBar & Phelps, 1998; Phelps, LaBar, & Spencer, 1997; Revelle & Loftus, 1990). Percentage of words recalled during the study phase averaged 8.8% ($n = 60$), 24.4% ($n = 60$), and 46.0% ($n = 30$) for Study Blocks 2, 3, and 4, respectively. There was no indication that learning was better or faster for words loading on the evaluative dimension. For example, for the group receiving four study blocks, percentage of recall for words loaded on the evaluation, potency, and activity dimensions averaged 10.2, 11.9, and 8.5, respectively, on the second study block, $F(2, 58) = 1.36$, $MSE = 0.006$, *ns*. The improvement in recall from the second to the fourth trial amounted to 30.4, 40.6, and 36.2, respectively, $F(2, 58) = 6.99$, $MSE = 0.011$, $p < .01$, so that, if anything, it was smaller for the evaluative dimension.

The results for the test phase also failed to yield evidence that words loading on the evaluative dimension are better recalled. These results indicated significant effects for number of study blocks, $F(1, 56) = 10.40$, $p < .01$, and for retention interval, $F(1, 56) = 50.32$, $p < .0001$. Across all words and participants, recall dropped sharply over the 1-week interval from 49.7% to only 19.2%, $F(1, 58) = 43.95$, $p < .0001$.

Commission errors. Access to the semantic attributes of words can be inferred from the nature of the commission errors made during recall (see Experiment 3 in Koriat, 1993). Because of the relatively small frequency of such errors, the analyses were carried out for all participants combined, pooling data across the study and test phases. Participants made a commission error in 10% of the cases, and 74% of these errors were within-list commissions and hence could be readily classified in terms of their semantic-differential ratings. We examined two questions. First, do commission errors tend to be of the same polarity on the pertinent dimension as the correct target? Second, is the tendency to make same-polarity commission errors evident only or mostly for the evaluative dimension?

[2] Invented words were used because people can sometimes guess the semantic attributes of words in noncognate natural languages and can even monitor the accuracy of their guesses (Koriat, 1975).

Table 1 presents the frequency of commission errors as a function of the dimension and pole of the nonrecalled target word and those of the commission response. With two exceptions, each of the frequencies appearing along the diagonal is the highest of any of the frequencies in the respective rows and columns. There were 208 errors in which the dimension and pole of the commission response matched that of the nonrecalled target, when the frequency expected on the basis of the marginals was only 127.7. This match was not confined to the evaluative dimension: The observed frequencies of pole-consistent errors for the evaluation, potency, and activity dimensions were 90, 63, and 55, respectively; whereas the expected frequencies were 51.0, 35.6, and 41.1, respectively. The respective ratios of observed-to-expected frequencies were 1.76, 1.77, and 1.34.

Traditional statistical procedures cannot be applied to the results presented in Table 1 because of the "fragmentary data problem" noted in connection with TOT studies (see A. S. Brown, 1991): Each cell is based on a different combination of participants and items. Therefore, we used the following procedure to evaluate the extent to which commission errors tend to be from the same pole as the nonrecalled target: For both target words and commission words, words classified as "bad," "weak," and "passive" were assigned a score of 1, and those classified as "good," "strong," and "active" were assigned a score of 2. For each participant, we calculated separately the mean score of the commission words for targets scoring 1 and 2 on each dimension. These means were then averaged across all participants for whom both means were available, and the group means are presented in Table 2 for each of the three dimensions.

Across the three dimensions, the average commission scores for targets assigned 1 and 2 were 1.42 and 1.63 (based on 54 participants), respectively, $F(1, 53) = 22.46$, $p < .0001$. We found this pattern of consistent polarities between commission errors and targets for each of the three dimensions, but it was significant only for the evaluative and potency dimensions, $F(1, 38) = 12.66$, $p < .001$; and, $F(1, 34) = 6.82$, $p < .05$, respectively; but not for the activity dimension ($n = 32$; $F < 1$). There were only 16 participants who provided at least one intralist commission error for each of the two poles of each dimension. A two-way analysis of variance (ANOVA), Dimension × Polarity (1 vs. 2), for these participants yielded $F(1, 15) = 3.75$, $p < .08$, for polarity; but $F < 1$ for both dimension and the interaction.

Table 2
Mean and Standard Error of the Mean of Commission Scores for Targets Belonging to Different Poles of Each Dimension in Experiment 1

	Dimension					
	Evaluation		Potency		Activity	
Target polarity	1	2	1	2	1	2
M	1.36	1.61	1.42	1.64	1.55	1.64
SEM	0.057	0.053	0.069	0.065	0.066	0.061
n	39		35		32	

Note. Targets scored 1 and 2 on each dimension.

The number of intralist commission errors in the test phase was too small to permit an analysis of the effects of retention interval. It should be noted, however, that for the immediate condition, the average commission scores for targets assigned 1 and 2 were 1.34 and 1.64, respectively, based on 19 participants for whom both means were available, $F(1, 18) = 7.42$, $p < .05$. The respective means for the delayed condition were 1.50 and 1.52 ($n = 19$), respectively ($F < 1$). A two-way Polarity × Retention interval yielded $F(1, 36) = 3.34$, $p < .08$, for polarity; $F < 1$ for retention interval, and, $F(1, 36) = 2.36$, $p < .11$, for the interaction. Although the interaction was not significant, the means suggest that access to attribute information, as disclosed by the type of commission errors made, decays over the 1-week retention interval.

Attribute identification. We obtained ratings on the semantic-differential dimensions on words for which the participant either failed to provide any answer or made a commission error (65.5% of the words, on average). Recall that these ratings were made on a 4-point scale so that *sure-bad*, *sure-weak*, and *sure-passive* received a score of 1, and *sure-good, sure-strong*, and *sure-active*

Table 1
Frequency Distribution of Commission Errors as a Function of the Dimension and Pole of the Nonrecalled Target Word and Those of the Commission Response in Experiment 1

Pole of nonrecalled target	Pole of Commission Response						
	Good	Bad	Strong	Weak	Active	Passive	Total
Good	38	18	14	15	22	22	129
Bad	19	52	20	22	17	26	156
Strong	19	20	29	15	30	7	120
Weak	20	11	11	34	15	15	106
Active	23	13	22	21	41	13	133
Passive	28	14	12	27	24	14	119
Total	147	128	108	134	149	97	763

Note. The boxes along the diagonal include the distribution of commission errors for cases in which the dimension of the commission error corresponded to that of the nonrecalled word.

received a score of 4. Table 3 presents the distribution of participants' ratings as a function of the pole of the target on the respective dimension.

The results suggest that when participants failed to retrieve the target, they were successful in judging its polarity on the respective dimension. To evaluate the significance of this trend, we averaged the ratings provided by each participant for words representing each of the poles of each dimension. The means of these means are presented in Figure 1 for the immediate (Figure 1A) and delayed conditions (Figure 1B), using in each case only participants for whom all six means were available. A Dimension × Pole ANOVA, collapsing data across both conditions (3 participants who did not have ratings on both poles of each dimension were eliminated from this analysis), yielded $F(1, 56) = 32.62$, $p < .0001$, for pole; and $F < 1$ for both dimension and the interaction. Separate one-way ANOVAs for each of the dimensions yielded $F(1, 56) = 7.94$, $p < .01$, for the evaluative dimension; $F(1, 56) = 24.11$, $p < .0001$, for the potency dimension; and, $F(1, 56) = 5.91$, $p < .05$, for the activity dimension. Thus, participants' judgments about the attributes of irretrievable words were significantly accurate for each of the three dimensions of the semantic differential.

How was the success of attribute identification affected by retention interval? A three-way ANOVA, Dimension × Pole × Retention interval, yielded $F(1, 55) = 35.99$, $p < .0001$, for pole and no other significant effect. We calculated a *discrimination score,* defined as the mean difference between the two poles (with higher scores representing more accurate judgments), for each of the three dimensions for each of the retention intervals. The mean discrimination score across the three dimensions was .29 for the immediate condition and .28 for the delayed condition ($F < 1$), suggesting little loss of attribute information over a 1-week interval. Separate one-way ANOVAs for the effect of pole for each retention interval yielded $F(1, 28) = 12.00$, $p < .005$, for the immediate test; and, $F(1, 29) = 26.78$, $p < .0001$, for the delayed test. In summary, the results suggest that explicit access to partial information exhibits little or no decay over a 1-week interval. It should be recalled that the drop in item recall during this period was quite substantial, from 49.7% to 19.2%.

Discussion

The results from the analysis of both commission errors and attribute-identification responses indicated that when recall of a

Table 3
Frequency Distribution of Participants' Ratings as a Function of the Pole of the Target on the Respective Dimension in Experiment 1

	Attribute judgments			
Target	1	2	3	4
Bad	71	98	125	60
Good	39	86	116	67
Weak	53	94	81	51
Strong	28	77	117	82
Passive	66	97	89	57
Active	63	74	114	82

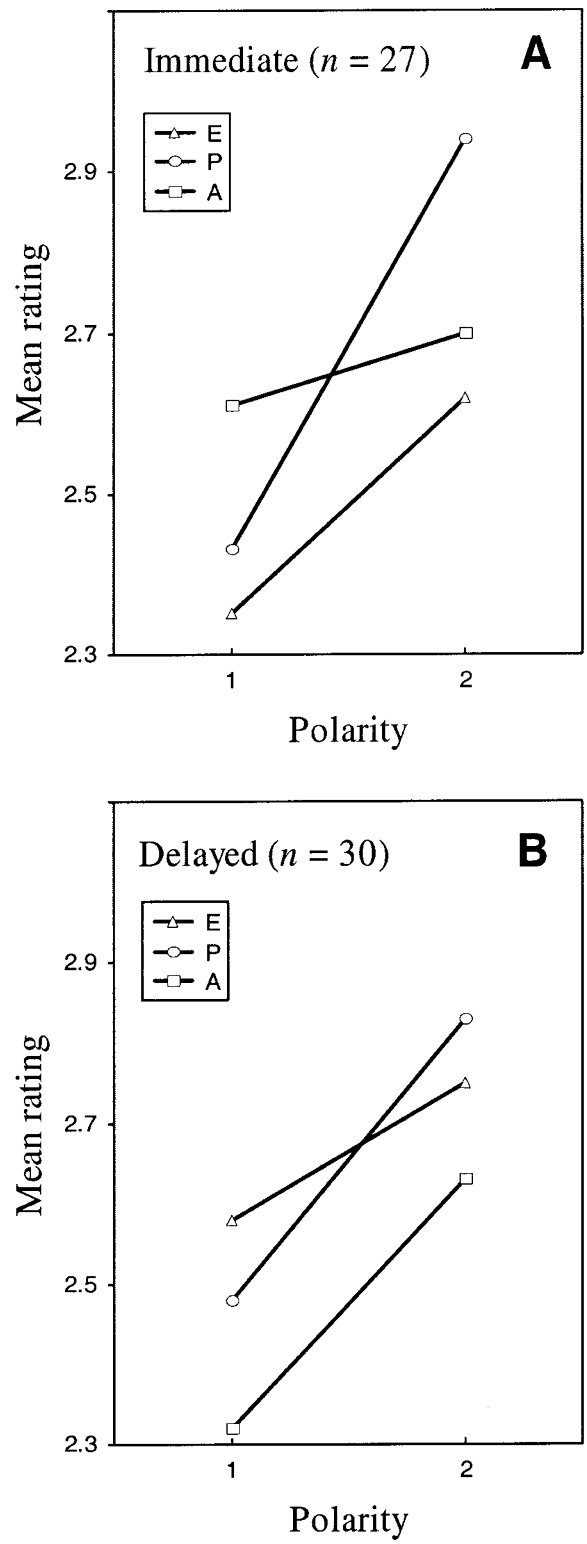

Figure 1. Mean ratings of words representing each of the poles of the evaluation (E), potency (P), and activity (A) dimensions for the immediate condition (Panel A) and the delayed condition (Panel B) in Experiment 1.

word fails, some partial semantic information about it may still be accessible. That information is not confined to the evaluative–emotional dimension, but occurs for the potency and activity dimensions as well.

The results also supported the hypothesis regarding the different forgetting rates for item recall and attribute identification. Whereas item recall dropped over the 1-week interval, from about 50% to about 20%, explicit access to the attribute polarity of the nonrecalled words evidenced virtually no change. Note that the assessment of partial information was based on a different pool of items in each of the retention intervals, but it is unlikely that this could be responsible for the observed difference in rate of information loss (see the General Discussion section). The results for commission errors were less conclusive on this matter.

Experiment 2

As noted in the introduction (and as suggested by the restaurant example), retrieval blocks are often accompanied by metacognitive experiences concerning the accuracy of the partial information retrieved. Experiment 2 focused on the phenomenal experiences associated with partial recall in comparison with item recall and examined the extent to which they are diagnostic of the accuracy of the response.

The experiment tested three hypotheses. The first is that attribute recall should be associated to a lesser extent with a recollective experience than item recall. Hence, using the know-remember distinction proposed by Tulving (1985), we expect partial recall to be associated with a greater frequency of know than of remember responses in comparison with item recall. A "guess" category was included, as is the recommended practice in current research on the know-remember distinction (Gardiner & Richardson-Klavehn, 2000).

The second hypothesis is that the phenomenal experience associated with an answer is diagnostic of the accuracy of that answer. Many observations indicate that the phenomenological state associated with blocked recall is diagnostic of the accuracy of the partial information retrieved and of the success of recalling or recognizing the target in a subsequent test (e.g., Koriat, 1993; Kozlowski, 1977; Schacter & Worling, 1985; Schwartz, Travis, Castro, & Smith, 2000). Schwartz et al., for example, observed that the likelihood of TOT resolution (i.e., subsequent recall of the target) could be predicted from subjective ratings of TOT intensity, the degree of emotionality accompanying TOT, and the feeling of recall imminence. In Experiment 2, we examined the hypothesis that item and attribute recalls are most likely to be correct when they are associated with a remember phenomenal state and least likely to be correct when they are associated with a "guess" state. However, assuming that source monitoring is more difficult for partial recall than for full recall, then we might expect the diagnostic validity of phenomenal experience to be lower for attribute identification than for item recall.

The third hypothesis is that when participants are forced to provide attribute identification, then that identification may prove accurate even when they initially deny any knowledge of the solicited target. Indeed, several previous studies have indicated that participants are accurate in making semantic judgments about a rare word even when they deny any knowledge of that word and are reluctant to bet on its meaning (Durso & Shore, 1991; Eysenck, 1979; Shore & Durso, 1990).

The procedure of Experiment 2 was similar to that of Experiment 1, except that (a) participants made know-remember judgments regarding their recall and attribute-identification responses; (b) the list was presented for three study blocks, and the test phase was always immediate; and (c) a final phase was included in which participants were forced to make attribute identifications for items to which they had made a don't know response on the preceding phase.

Method

Participants. Sixty Hebrew-speaking University of Haifa students (45 women and 15 men) participated in the experiment for course credit.

Materials and procedure. The list of stimuli was the same as that used in Experiment 1. The study phase was also the same, except that all participants were presented with three study blocks.

The test phase took place immediately after the study phase for all participants. In the first block of that phase, participants were presented with each Somali word for 8 s. When they responded with a Hebrew word, they were asked to indicate their state of awareness by clicking on one of three options: REMEMBER, KNOW, or GUESS. The instructions (in Hebrew) for making these judgments followed those presented by Gardiner and Richardson-Klavehn (2000). The next Somali word was then presented. However, when participants either failed to retrieve the word or responded "don't know," they were asked to guess the polarity of the nonrecalled word. The procedure was the same as in Experiment 1, except that participants made a two-choice response (rather than providing a response on a 4-point scale). Participants could either choose one of the poles by clicking on it with the mouse or they could say aloud "don't know" into the microphone and move on to the next Somali word. When participants made attribute identifications, they were asked to indicate their state of awareness by clicking on one of three options with the mouse: REMEMBER, KNOW, or GUESS. They were then presented with the next Somali word.

In a second test phase, only words for which participants had made a don't know response on the attribute-identification task were presented with the corresponding values (e.g., good–bad), and participants were now forced to make attribute identifications on these words.

Results

We first summarize the results replicating those obtained in Experiment 1.

Recall performance. Percentage of words recalled in Study Blocks 2 and 3 and in the test block averaged 12.2, 29.7, and 45.8, respectively. Across these blocks, recall averaged 28.0, 32.1, and 27.6, respectively, for the evaluation, potency, and activity dimensions. Thus, as in the previous experiment, there was no indication of better recall for words loading on the evaluative dimension.

Commission errors. Pooling data across the study and test phases, participants made a commission error in 11.7% of the cases, and 68.7% of these errors were within-list commissions. The distribution of the within-list commission errors revealed that there were 196 errors in which the dimension and pole of the commission response matched that of the nonrecalled target, compared with an expected frequency of 115.48. There were no systematic differences between the evaluation, potency, and activity dimensions in this respect.

Attribute identification. When participants either failed to produce any answer or responded "don't know" in the item-recall task

(43.6% of the trials), their attribute identification was correct in 56.2% of the cases, $t(58) = 2.77$, $p < .01$. A one-way ANOVA comparing attribute identification for the three dimensions (using 44 participants for whom attribute-identification scores were available for all three dimensions) yielded $F < 1$. Percentage of accuracy for these participants averaged 55.9, 55.9, and 51.3, for the evaluation, potency, and activity dimensions, respectively.

The phenomenal quality of complete versus partial retrieval. We turn now to the hypotheses that were the main concern of Experiment 2, focusing first on the phenomenal states accompanying item recall and attribute recall. As can be seen in Table 4, attribute identifications were associated with a considerably larger proportion of guess responses than item recalls, $t(58) = 10.74$, $p < .0001$; despite the fact that in both cases, participants had the option to respond "don't know." Item recalls, in contrast, were associated with a larger proportion of remember responses than attribute identifications, $t(58) = 10.84$, $p < .0001$. For item recall, the percentage of remember judgments exceeded that of know judgments, $t(59) = 0.39$, $p < .0001$; whereas for attribute identification, there was a nonsignificant trend in the opposite direction.

The diagnostic value of state of awareness. We turn next to the second hypothesis, that is, state of awareness is diagnostic of the correctness of item recall and attribute identification. For item recall, the probability that the answer was correct averaged .91 for remember responses, .76 for know responses, and .32 for guess responses. A one-way ANOVA (with 42 participants for whom all means were available) yielded $F(2, 82) = 66.28$, $p < .0001$. We found a similar pattern for attribute identification: The respective probabilities were .71, .65, and .45, $F(2, 38) = 4.22$, $p < .05$, with 20 participants. Thus, state of awareness was diagnostic of the correctness of the answer for both item recall and attribute identification.

Note, however, that the relationship between state of awareness and memory accuracy was somewhat weaker for attribute identification than for item recall: A two-way ANOVA (with 17 participants) yielded $F(2, 32) = 29.01$, $p < .0001$ for the interaction. A post hoc analysis (Scheffé) on the results for item recall (with $n = 42$) indicated that all three states of awareness differed significantly from one another in terms of their associated probabilities of correct recall, whereas a similar analysis on the results for attribute identification (with $n = 20$) indicated that only the difference between remember and guess responses was significant.

We also examined the question of whether correct and incorrect answers tend to elicit different phenomenal states (see Table 5).

Table 4
Mean Percentage of Know and Remember Responses for Item and Attribute Recall Irrespective of the Correctness of the Answer in Experiment 2

	State of awareness					
	Remember		Know		Guess	
Memory measure	*M*	*n*	*M*	*n*	*M*	*n*
Item recall	63.2	60	23.5	60	13.4	60
Attribute identification	20.2	59	28.0	59	51.8	59

Table 5
Mean Percentage of Know and Remember Responses for Correct and Incorrect Item Recalls and for Correct and Incorrect Attribute Identifications Recalls (n = 51) in Experiment 2

	State of awareness		
Memory measure	Remember	Know	Guess
Item recall			
Correct	72.7	21.7	5.6
Incorrect	32.3	24.8	42.9
Attribute identification			
Correct	22.1	32.2	45.7
Incorrect	14.7	22.5	62.8

Indeed, correct answers elicited significantly fewer guess responses than incorrect answers: A Type of Recall (item recall vs. attribute identification) × Correctness ANOVA on the percentage of guess responses (using only 51 participants for whom all means were available) yielded $F(1, 50) = 82.53$, $p < .0001$ for correctness. Guess responses were more frequent for attribute identifications than for item recalls, as reported earlier, $F(1, 50) = 59.75$, $p < .0001$; but the interaction was also significant, $F(1, 50) = 11.75$, $p < .005$; suggesting a stronger discrimination in the proportion of guess responses between correct and incorrect item recalls than between correct and incorrect attribute identifications. However, incorrect answers were associated with a larger percentage of guess responses both in the case of item recall, $t(50) = 9.17$, $p < .0001$, as well as in the case of attribute identification, $t(50) = 3.92$, $p < .0005$.

Did correct and incorrect answers also differ in the pattern of know and remember responses? To examine this question, we ignored the "guess" category and calculated for each participant the percentage of remember responses out of the total number of know and remember responses together. For item recall, this percentage averaged 75.6% for correct responses and 51.6% for incorrect responses. The respective percentages for attribute identification were 37.7% and 35.2%. A two-way ANOVA, Type of Recall × Correctness, using only 34 participants for whom all 4 means were available, yielded $F(1, 33) = 21.11$, $p < .0001$, for type of recall; $F(1, 33) = 8.22$, $p < .01$, for correctness; and, $F(1, 33) = 6.90$, $p < .05$, for the interaction. Thus, when only remember and know responses are considered, correct responses tended to elicit a larger percentage of remember responses overall (56.7%) than incorrect responses (43.4%). However, this pattern was found only for item recall, $F(1, 33) = 11.25$, $p < .005$, and not for attribute identification ($F < 1$).

In summary, the state of awareness accompanying memory is diagnostic of memory accuracy for both item recall and attribute identification. However, whereas a guess response tended to single out the wrong answers in both cases, the know-remember distinction was diagnostic of the correctness of the answer only for item recall but not for attribute identification. Also, for item recall, correct answers elicited a different pattern of know-remember responses than incorrect answers, whereas this was not so for attribute identification.

Attribute identification for don't know responses. Were participants able to access partial information about an elusive target even when they denied any knowledge of that information? In the final phase of the experiment, participants were forced to make attribute identifications on items for which they had made a don't know response (an average of 39.2% across participants). There were 51 participants who had made a don't know response at least once, and for them mean correct identification in the final phase was 57.2, significantly better than chance, $t(50) = 2.22$, $p < .05$. In fact, attribute identification was no less accurate overall when participants initially made a don't know response (57.2%) than when they chose to make attribute identifications (56.2%). Thus, participants' guesses of the semantic attributes of an elusive word were correct better than chance even when they initially denied any knowledge of these attributes. This finding is somewhat surprising given that participants' attribute identifications were actually no better than chance when they were associated with a guess response.

Discussion

The results of Experiment 2 replicated the main findings of Experiment 1: Both the errors committed during recall and the attribute judgments revealed access to all three semantic dimensions of the nonrecalled words.

In addition, the results supported the hypotheses pertaining to phenomenal states and their diagnostic validity. First, attribute identification elicited fewer remember responses than item recall, consistent with the idea that partial recall is less likely to evoke a recollective experience that specifies contextual details. However, it elicited more guess responses than item recall despite the fact that participants were given the option to respond "don't know" in both the item-recall and the attribute-identification tasks. It is difficult to interpret these findings given that the accuracy of guess partial recalls was not better than chance. However, the overall pattern of results suggests that partial recall was more likely than full recall to be associated with a more fleeting phenomenal experience such as that of know or guess.

Note that not only was item recall sometimes associated with know responses but also attribute information was sometimes associated with remember responses. This pattern is consistent with recent evidence (e.g., Hicks, Marsh, & Ritschel, 2002; see also Conway & Dewhurst, 1995) suggesting that source monitoring is not an all-or-nothing process, and that even vague, partial information can support source monitoring.

The results pertaining to the diagnostic validity of phenomenal state were rather clear for item recall: The accuracy of the answer increased from guess to know to remember responses. The results for attribute identification, in contrast, were somewhat perplexing: Only the remember-guess distinction was diagnostic of memory accuracy, whereas the know-remember distinction was not. Furthermore, the tendency to volunteer the attribute of the elusive word was not diagnostic of memory accuracy because attribute identification was accurate better than chance even for items for which participants had initially declined to make such identification, and, in fact, no less accurate than identifications that participants chose to volunteer. This pattern of results is consistent with the claim (Durso & Shore, 1991) that access to partial information is based, in part, on implicit memory.

We had expected that attribute judgments associated with a guess response would also be accurate better than chance. This, however, was not the case. Thus, it would seem that the very decision to volunteer or withhold an attribute judgment is not diagnostic of the accuracy of that judgment, whereas when a judgment is volunteered, the phenomenal state associated with it is predictive of its accuracy. Could this pattern disclose a systematic difference between monitoring and control processes in the case of partial retrieval (see Koriat & Goldsmith, 1996), with control processes being less effective than monitoring processes? More work is necessary before we can reach definite conclusions on this matter.

General Discussion

This study had two objectives: first, to examine the kind of semantic information that people access about a word that they cannot recall, and second, to gain some insight into the characteristics of the process underlying partial recall. We discuss the evidence pertaining to each of these objectives in turn.

Accessing the Semantic Attributes of Inaccessible Words

It is surprising how little research has been conducted on the question of what semantic features of a word a person can access in the face of failing to recall the word itself. Research on this question is expedient because it can shed light on both the representation of semantic information in memory (e.g., Collins & Loftus, 1975) and the processes by which stored information comes to mind (R. W. Brown & McNeill, 1966; Koriat & Lieblich, 1974), or finds its way into behavior (Wheeler & Petty, 2001).

The results of the present study clearly indicate that information about the semantic attributes of a word may be accessible even when the word itself cannot be recalled. These results are consistent with the view that memory and forgetting are not all-or-nothing processes (Brainerd et al., 1990; R. W. Brown & McNeill, 1966). Rather, concepts and episodes are represented in memory as a bundle of attributes that are bound together to different degrees, and that can be forgotten or remembered with relative independence from one another (Schacter et al., 1998; Underwood, 1969). In fact, it has been proposed that even when retrieval is successful, it rarely involves the recollection of a coherent record in its entirety but more often entails the assembly and fusion of several related fragments (P. T. Smith, 2000). Presumably, then, when some of the features of a word are lost, or when the cohesion between these features weakens, item recall may fail and yet access to some of the individual semantic features of the word survives.

Our results may also be seen to support feature-based models of semantic memory according to which the meaning of words is represented as a list of features (E. E. Smith, Shoben, & Rips, 1974). In terms of E. E. Smith et al.'s model, the features of the words whose recall was tested in the present study constitute, perhaps, the defining or essential features of some of the words (e.g., *bad* for *danger*, *light* for *feather*). The focus on such features was dictated, in part, by the methodological constraint of having to equate the words in terms of their degree of polarity on their respective dimensions. Thus, perhaps these are the features that first come to mind when complete recall fails, or that constrain the type of commission errors that participants produce.

Previous studies on the retrieval of partial semantic information have focused on the evaluative-emotional dimension of words, perhaps reflecting the conviction that this dimension is the most critical attribute of words and objects (see Strack & Deutsch, 2002). Indeed, Zajonc's (1980, 1984) affective primacy thesis may be taken to imply that perhaps only the emotional qualities of a word can be identified when the full word evades recall: "We are reminded of an interpersonal conflict of long ago. The cause of the conflict, the positions taken, the matter at issue, who said what, may have all been forgotten, and yet the affect that was present during the incident may be readily retrieved" (1980, p. 159).

However, to the best of our knowledge, the affective primacy hypothesis has been tested and supported only with regard to the reactions to external input. The present study, in contrast, failed to find evidence for affective primacy with regard to the retrieval of information from memory. Rather, access to an elusive memory word appears to be multifaceted, occurring along several semantic dimensions. This was true for commission errors as well as for explicit attribute identification.

The failure to find evidence for the affective primacy hypothesis either in commission errors or in attribute judgments should come as a surprise in view of many findings that have been taken to suggest that affective information is processed faster than purely cognitive information, and that this processing can occur automatically and below full awareness (e.g., Bargh, 1997; Duckworth et al., 2002; Murphy & Zajonc, 1993). Thus, in the same way that people have privileged access to the emotional tone of a suboptimally presented word that they fail to identify, we might have expected the same to be true of an elusive word that people are trying to retrieve from memory. What is the explanation of this discrepancy? Operationally, there are several procedural differences between the task of judging the semantic attributes of a subliminal word and that of judging the attributes of a previously studied word in response to a cue, and each of these may contribute to the different pattern of results obtained. However, we should also consider the possibility that our results disclose a more fundamental difference between perception and memory, that is, between accessing information from without and accessing information from within. From a functional point of view, there is benefit in the ability to rapidly classify external stimuli in terms of their implications for approach–avoidance behavior (Bargh, 1997; Chen & Bargh, 1999; Lang, Bradley, & Cuthbert, 1990; Neumann & Strack, 2000). Certainly, under emergency conditions, the ability to quickly detect threatening stimuli should be critical for adaptation and survival. However, perhaps the urgency of identifying threatening memories is not as severe as that of detecting threats that originate from the outside world. Needless to say, these remarks are speculative and more work is needed to clarify this issue.

The Characteristics of Partial Retrieval

The second aim of this study was to test two hypothesized differences between the retrieval of a memory entry and the retrieval of only partial information about it. Assuming that a failure to recall a word occurs when some of the features of the word are lost or when the cohesion between these features loosens, then we may expect attribute memory to outlast item memory. Furthermore, attribute recall should be associated to a lesser extent with recollective experience than word recall.

Both of these hypotheses received some support. With regard to the first hypothesis, Experiment 1 indicated that whereas item recall dropped from 49.7% to 19.2% over the 1-week interval, correct attribute identification evidenced practically no drop at all over that interval. This pattern was replicated in another unpublished experiment that was quite similar in procedure to Experiment 1. In that experiment, item recall dropped from 38.5% to 12.5% over a 1-week interval, whereas explicit access to partial information did not decline at all over that interval.[3]

A methodological problem concerning this comparison, however, must be addressed. The item-recall scores and the attribute-identification scores are based on different sets of items, because for each participant, attribute identification was measured on those items for which he or she failed to retrieve the correct answer. A similar problem exists in feeling-of-knowing (FOK) studies (see Schwartz & Metcalfe, 1994). Is it possible, then, that the contrast between item and attribute recall actually reflects a between-item difference? This seems to be unlikely. In fact, Koriat (1995) observed that items that yielded a high proportion of complete recall across participants also yielded more partial information when recall of the target failed. Koriat and Lieblich (1977) also reported a similar pattern: Word definitions that produced a high proportion of correct recalls also precipitated a relatively high proportion of "TOT-Intended" and "TOT-Got it-Correct" states when initial recall failed. This correlational pattern makes it all the more impressive that partial semantic access is preserved over a relatively long retention interval for inaccessible words.

The second hypothesis, which concerns the subjective experience associated with partial and full recall, also received support: Item recalls were predominantly associated with remember responses, whereas attribute judgments were predominantly associated with know and guess responses, suggesting that source monitoring is more difficult for partial recall than for complete recall. The results also suggested that participants are less successful in monitoring the accuracy of their attribute identifications than in monitoring the accuracy of item recall. First, the know-remember distinction was diagnostic of memory accuracy for item recall but not for attribute identification. Second, attribute judgments that were initially withheld turned out to be no less accurate than those that were volunteered. A similar pattern was observed by Durso and Shore (1991). Although participants were not willing to bet on the correct meaning of words that they had classified as nonwords, they performed with above chance success when forced to do so.

What are the implications of these results? We believe that they are consistent with the general view that both the access to partial information about an elusive memory target and the metacognitive feelings associated with it provide some clues regarding the way in which stored information about which one is not fully aware

[3] A complete description of the unpublished experiment is available from Asher Koriat upon request.

affects subjective experience and performance (see Koriat, 2000; Koriat & Levy-Sadot, 1999). A similar assumption underlies the analysis of Dorfman, Shames, and Kihlstrom (1996), who likened the process underlying FOK and TOT to that underlying the intuitive feeling that people sometimes have about the solution of a problem before they reach it.

References

Anisfeld, M., & Knapp, M. E. (1968). Association synonymity and directionality in false recognition. *Journal of Experimental Psychology, 77,* 171–179.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer Jr. (Ed.), *The automaticity of everyday life: Advances in social cognition* (Vol. 10, pp. 1–61). Mahwah, NJ: Erlbaum.

Bargh, J. A., Litt, J., Pratto, F., & Spielman, L. A. (1989). On the preconscious evaluation of social stimuli. In A. F. Bennet & K. M. McConckey (Eds.), *Cognition in individual and social contexts: Proceedings of the XXIV International Congress of Psychology* (Vol. 3, pp. 357–370). Amsterdam: Elsevier/North-Holland.

Bargh, J. A., Raymond, P., & Chaiken, S. (1995). *The automatic evaluation effect: Does it hold for the other major dimensions of semantic meaning as well*? Unpublished manuscript, New York University.

Belli, R. F., Windschitl, P. D., McCarthy, T. T., & Winfrey, S. E. (1992). Detecting memory impairment with a modified test procedure: Manipulating retention interval with centrally presented event items. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 356–367.

Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review, 100,* 42–67.

Brainerd, C. J., Reyna, V. F., Howe, M. L., & Kingma, J. (1990). The development of forgetting and reminiscence. *Monographs of the Society for Research in Child Development, 55*(Nos. 3–4, Serial No. 222), 1–111.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109,* 204–223.

Brown, R. W., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5,* 325–337.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25,* 215–224.

Cohen, G., Stanhope, N., & Conway, M. A. (1992). Age differences in the retention of knowledge by young and elderly students. *British Journal of Developmental Psychology, 10,* 153–164.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 407–428.

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

Dorfman, J., & Mandler, G. (1994). Implicit and explicit forgetting: When is gist remembered? *The Quarterly Journal of Experimental Psychology, 47A,* 651–672.

Dorfman, J., Shames, V. A., & Kihlstrom, J. F. (1996). Intuition, incubation, and insight: Implicit cognition in problem solving. In G. Underwood (Ed.), *Implicit cognition* (pp. 257–296). Oxford, England: Oxford University Press.

Duckworth, K. L., Bargh, J. A., Garcia, M., & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science, 13,* 513–519.

Durso, F. T., & Shore, W. J. (1991). Partial knowledge of word meanings. *Journal of Experimental Psychology*: *General, 120,* 190–202.

Eysenck, M. W. (1979). The feeling of knowing a word's meaning. *British Journal of Psychology, 70,* 243–251.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitude. *Journal of Personality and Social Psychology, 50,* 229–238.

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

Heise, D. R. (1965). Semantic differential profiles for 1,000 most frequent English words. *Psychological Monographs: General and Applied, 79*(8, Whole No. 601), 1–12.

Hicks, J. L., Marsh, R. L., & Ritschel, L. (2002). The role of recollection and partial information in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 503–508.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language, 29,* 133–159.

Koriat, A. (1975). Phonetic symbolism and the feeling of knowing. *Memory & Cognition, 3,* 545–548.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100,* 609–639.

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

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition, 9,* 149–171.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,* 490–517.

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 483–502). New York: Guilford Press.

Koriat, A., & Lieblich, I. (1974). What does a person in a "TOT" state know that a person in a "don't know" state doesn't know? *Memory & Cognition, 2,* 647–655.

Koriat, A., & Lieblich, I. (1977). A study of memory pointers. *Acta Psychologica, 41,* 151–164.

Kozlowski, L. (1977). Effects of distorted auditory and rhyming cues on retrieval of tip-of-the-tongue words by poets and nonpoets. *Memory & Cognition, 5,* 477–481.

LaBar, K. S., & Phelps, E. A. (1998). Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science, 9,* 490–493.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97,* 377–395.

Lovelace, E. (1987). Attributes that come to mind in the TOT state. *Bulletin of the Psychonomic Society, 25,* 370–372.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64,* 723–739.

Nelson, T. O., Fehling, M. R., & Moore-Glascock, J. (1979). The nature of semantic savings for items forgotten from long-term memory. *Journal of Experimental Psychology: General, 108,* 225–250.

Neumann, R., & Strack, F. (2000). Approach and avoidance: The influence of proprioceptive and exteroceptive cues on encoding of affective information. *Journal of Personality and Social Psychology, 79,* 39–48.

Niedenthal, P. M. (1990). Implicit perception of affective information. *Journal of Experimental Social Psychology, 26,* 505–527.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin, 49,* 197–237.

Phelps, E. A., LaBar, K. S., & Spencer, D. D. (1997). Memory for

emotional words following unilateral temporal lobectomy. *Brain and Cognition, 35,* 85–109.
Revelle, W., & Loftus, D. A. (1990). Individual differences and arousal: Implications for the study of mood and memory. *Cognition and Emotion, 4,* 209–237.
Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814.
Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology, 49,* 289–318.
Schacter, D. L., & Worling, J. R. (1985). Attribute information and the feeling of knowing. *Canadian Journal of Psychology, 39,* 467–475.
Schwartz, B. L. (2002). *Tip-of-the-tongue states: Phenomenology, mechanism, and lexical retrieval.* Mahwah, NJ: Erlbaum.
Schwartz, B. L., & Metcalfe, J. (1994). Methodological problems and pitfalls in the study of human metacognition. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 115–135). Cambridge, MA: MIT Press.
Schwartz, B. L., Travis, D. M., Castro, A. M., & Smith, S. M. (2000). The phenomenology of real and illusory tip-of-the-tongue states. *Memory & Cognition, 28,* 18–27.
Shore, W. J., & Durso, F. T. (1990). Partial knowledge in vocabulary acquisition: General constraints and specific detail. *Journal of Educational Psychology, 82,* 315–318.
Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81,* 214–241.
Smith, P. T. (2000). A jigsaw puzzle theory of memory. *Memory, 8,* 245–264.
Strack, F., & Deutsch, R. (2002). *Reflective and impulsive determinants of social behavior.* Manuscript submitted for publication.
Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26,* 1–12.
Underwood, B. J. (1969). Attributes of memory. *Psychological Review, 76,* 559–573.
Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin, 127,* 797–826.
Wickens, D. D., & Cermak, L. S. (1967). Transfer effects of synonyms and antonyms in mixed and unmixed lists. *Journal of Verbal Learning and Verbal Behavior, 6,* 832–839.
Yarmey, A. D. (1973). I recognize your face but I can't remember your name: Further evidence on the tip-of-the tongue phenomenon. *Memory & Cognition, 1,* 287–290.
Yavuz, H. S., & Bousfield, W. A. (1959). Recall of connotative meaning. *Psychological Reports, 5,* 319–320.
Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175.
Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist, 39,* 117–123.

Appendix

The 48 Words (Translated From Hebrew) Representing the Two Poles of the Three Dimensions of the Semantic Differential

Evaluation		Potency		Activity	
Good	Bad	Strong	Weak	Active	Passive
Happy	(to) *Fall*	*Army*	*Baby*	*Fast*	*Nothing*
Talent	*Danger*	*Rigid*	*Feather*	(to) *Grow*	(to) *Sit*
Pleasure	*Enemy*	*Law*	*Elderly*	*Movement*	(to) *Wait*
Fresh	*Failure*	*Steel*	*Chick*	*Stream*	*Boredom*
Friend	(to) *Kill*	*Mountain*	*Hair*	(to) *Run*	*Tired*
Health	(to) *Lose*	*Weight*	*Grain*	*Change*	*Pumpkin*
Quality	*Accident*	*Feast*	(light) *Bulb*	*River*	*Paper*
Peace	*Pain*	*Axe*	*Twig*	*Forward*	*Sofa*

Received July 10, 2002
Revision received March 25, 2003
Accepted March 26, 2003 ■