Recalling taboo and nontaboo words

TIMOTHY JAY
Massachusetts College of Liberal Arts

CATHERINE CALDWELL-HARRIS
Boston University

KRISTA KING
New York University

People remember emotional and taboo words better than neutral words. It is well known that words that are processed at a deep (i.e., semantic) level are recalled better than words processed at a shallow (i.e., purely visual) level. To determine how depth of processing influences recall of emotional and taboo words, a levels of processing paradigm was used. Whether this effect holds for emotional and taboo words has not been previously investigated. Two experiments demonstrated that taboo and emotional words benefit less from deep processing than do neutral words. This is consistent with the proposal that memories for taboo and emotional words are a function of the arousal level they evoke, even under shallow encoding conditions. Recall was higher for taboo words, even when taboo words were cued to be recalled after neutral and emotional words. The superiority of taboo word recall is consistent with cognitive neuroscience and brain imaging research.

Examples from everyday life provide ample evidence that information and events associated with strong emotions are remembered better than experiences that lack emotional depth. We remember when and where we learned about 9/11 attacks on the World Trade Center and Pentagon but not the clothing we wore the day before. People also feel more confident about the accuracy of their memories they view as emotionally charged, whether they are more accurate or not (Sharot, Delgado, & Phelps, 2004). Laboratory experiments with emotional material have confirmed everyday impressions regarding the effect of emotion on memory (see Rapaport, 1942, for a survey of early experiments on emotions and memory). Research on flashbulb memories indicates that people retain vivid, detailed memories of emotional events (Brown & Kulik, 1977).

Our research focuses on emotionality and memory for taboo words. Taboo words represent a class of emotionally arousing references with respect to body products, body parts, sexual acts, ethnic or racial insults, profanity, vulgarity, slang, and scatology (Jay, 1992, 2000). Emotionally

AMERICAN JOURNAL OF PSYCHOLOGY
Spring 2008, Vol. 121, No. 1, pp. 83–103

arousing words are remembered better than nonarousing words, and taboo words show the most exaggerated version of this effect (e.g., Kensinger & Corkin, 2003; LaBar & Phelps, 1998; MacKay & Ahmetzanov, 2005; MacKay et al., 2004). One compelling reason for superior recall of taboo words is based on their emotional qualities. Taboo words have uniquely strong connotative meanings; in fact, their primary meaning is connotative, which is unusual relative to nontaboo words, which are more denotative. Research suggests that it is the emotional arousal attached to taboo words that makes them memorable (Anderson & Phelps, 2001a, 2001b; Kensinger & Corkin, 2003, 2004; Sharot & Phelps, 2004). Word arousal is the degree to which a word is calming (e.g., *water*) or arousing (e.g., *fuck*), which is correlated with physiological measures such as skin conductance response (SCR; Manning & Melchiori, 1974). Many taboo words have both negative valence (i.e., "bad" words) and arousal, which can be contrasted to words such as *sorrow,* which have negative valence but are not arousing.

Arousal and memory

Researchers have studied how taboo word encoding and arousal affect recall. For example, LaBar and Phelps (1998) and Maratos, Allan, and Rugg (2000) proposed that activation of the amygdala or prefrontal cortex influences memory. LaBar and Phelps (1998; also see Adolphs, Russell, & Tranel, 1999; Isenberg et al., 1999) studied the influence of emotional arousal on declarative memory, assuming that the underlying arousal–memory interactions were based on amygdala–hippocampal consolidation processes. Normal participants and temporal lobectomy patients rated emotionally arousing taboo words and neutral words on an arousal scale while their SCRs were monitored. Both arousal measures for both classes of participants indicated that taboo words were more arousing than neutral words at encoding; however, temporal lobectomy patients did not exhibit the attenuated forgetting that normal participants did. Only the normal participants showed an increase in memory for taboo words, suggesting a role for medial temporal lobe structures in memory consolidation over time for taboo words. Normal participants, but not patients, produced arousal-mediated memories, implicating a central role for the amygdala and medial temporal lobe areas in taboo word processing.

More recently, Kensinger and Corkin (2004) investigated brain structures and pathways involved in processing emotional words using functional magnetic resonance imaging and behavioral methods. They found separable cognitive processes and neural pathways for arousing words and nonarousing words. Their research suggests a dual-route model for word processing. Arousing taboo words are associated with processing in the amygdalar–hippocampal pathway; however, nonarousing words rely

on controlled processing in the hippocampal–prefrontal pathway. The prefrontal cortex uses controlled cognitive processes such as elaboration to enhance memories for negative nonarousing words. If a divided attention task overburdens attentional resources used for elaboration, then negative nonarousing words no longer evidence memory enhancement. In contrast, arousing words, which use the amygdalar network, do not suffer from a divided attention task, reflecting the highly automatic effects of emotional words on memory. This study and others (Anderson, 2005; Anderson & Phelps, 2001a, 2001b; MacKay & Ahmetzanov, 2005) indicate that the arousing properties of taboo words produce vivid memories without elaborative processing, which may be sufficient to override the burden of limited attentional resources. On the other hand, nonarousing words rely on elaborative processing for better memories.

LOP and memory

Another explanation for better memory for taboo words relative to nontaboo words is that taboo words may be encoded more effectively than nontaboo words. A levels of processing (LOP) model proposed by Craik and Lockhart (1972) viewed memory for verbal material as a function of encoding. The level of encoding that each word receives determines its memorability. A shallow level of processing is one in which only superficial or physical aspects are encoded. A deeper level of processing takes more time and effort to activate the semantic meaning of the stimulus. For example, when words are the stimuli, shallow processing involves scanning the word for its physical characteristics (uppercase or lowercase font). Shallow processing results in poor recall. Words that receive a deep level of processing, such as making a meaningful or semantic connection to the word, would persist longer. A semantic decision about a word (e.g., does it fit into the sentence frame "A _____ is part of the central nervous system?") would lead to a deep level of processing and produce a memory trace that remained active longer than those for words given shallow encoding. Positron emission tomography images confirm significantly more activation in the left prefrontal cortex in a semantic task than in a shallow task (Kapur et al., 1994). It could be the case that taboo words, regardless of orienting task, attract a deeper level of processing than nontaboo words and therefore are remembered better than nontaboo words. This deep level of initial encoding effect has been demonstrated with picture stimuli (Grady, McIntosh, Rajah, & Craik, 1998; Kern, Libkuman, & Otani, 2002; Kern, Libkuman, Otani, & Holmes, 2005).

Clearly the relationship between encoding processes and retrieval processes influences final recall rates, as demonstrated in classic studies of transfer-appropriate learning and encoding specificity (e.g., Bower, 1981; Morris, Bransford, & Franks, 1977; Tulving & Thompson, 1973). Gener-

ally, the more the retrieval environment matches the original encoding environment, the better the recall. The encoding specificity principle (Tulving & Thompson, 1973) explains the encoding–retrieval relationship as follows: The information that is encoded about a stimulus will be beneficial if present during the retrieval process. How does emotion affect this process? According to Bower (1981), the emotional context of an event gets encoded along with other semantic and episodic information. Bower examined processes that influence both encoding and retrieval stages of memory. Encoded emotional information about a word will prove most helpful when episodic information at recall is weak. If the recall environment is rich with retrieval cues, emotional information will be less useful. The emotional arousal associated with taboo words should be helpful during retrieval when other cues are weak.

For memory for emotional words, Reber, Perrig, Flammer, and Walter (1994) demonstrated that with shallow processing, emotional words are recalled better than neutral words. Reber et al. concluded that the emotional qualities of words provide a memory-enhancing effect only when they are encoded at a shallow level, in accordance with Bower's (1981) network theory of emotion. Bower developed a semantic network approach to emotional memory in which each distinct emotion has a specific node or unit in memory that connects to pertinent episodic information. Episodic information encapsulating an emotional event (who, what, where, when) is connected to related emotion nodes so that memory for an emotional event contains both semantic and affective information. Bower's network theory of emotion predicts that emotional valence influences recall only when minimal retrieval cues are available, as in shallow processing. When a rich pool of retrieval cues is available, as in deep processing, the emotion advantage is attenuated.

The effect of emotional arousal on memory persists after encoding has taken place. The effect of arousal on memory also may occur after encoding during a retention interval (Kleinsmith & Kaplan, 1963; Sharot & Phelps, 2004). In these cases memory for arousing materials is better after a delay than it is immediately after encoding. The retention and retrieval arguments are supported by the finding that memory for neutral stimuli decreases over time, whereas recall of arousing words improves or remains the same over time. Focusing on encoding or retrieval strategies rather than arousal to explain the link between emotion and memory raises an interesting question: Which is more important for memory, a word's arousal level or its level of encoding?

The current study

Both word arousal and LOP (Jacoby & Dallas, 1981) were varied to investigate competing explanations for why taboo words are remembered

better than nontaboo words. If word arousal is a stronger predictor of recall than LOP, then taboo words are expected to have superior recall regardless of deep or shallow processing. Little work has been done to manipulate level of encoding with taboo words.

Sharot and Phelps (2004) considered LOP to explain why arousing words were remembered better than neutral words in the periphery when attention was focused in a central screen location. They explained that although the attentional resources that are allocated to arousing and neutral words are similar, arousing words may profit from a more elaborate level of processing than neutral words, increasing long-term retention. It is unclear what is meant by "elaborative" processing with their procedure. Participants made frequency judgments about words presented for 250 ms in the central field and later made recognition judgments about the peripheral words. Traditionally, elaborative encoding engages semantic processing; however, the denotative meanings of peripheral words are not necessarily activated here. Recognition judgments can be made on the basis of shallow processing alone. Perhaps some form of connotative meaning that mediates arousal was activated, but the actual processing level of peripheral words is unknown because their level of processing was not manipulated directly. Sharot and Phelps ruled out the LOP hypothesis because it would have predicted better memory for arousing words than neutral words at immediate and delayed recognition intervals. They obtained better recall for arousing words only after a delay. It could be the case that neither arousing nor nonarousing words were given elaborative processing, and their results reflect only the effects of arousal that takes place over time.

Importantly, Sharot and Phelps (2004) did not manipulate LOP directly with taboo words in the periphery. We propose that taboo words produce strong memories regardless of level of attention or encoding directed to them. Enhanced processing for taboo words regardless of LOP would be consistent with Kensinger and Corkin's (2004) dual-route theory of word processing, in which taboo words achieve vivid memories because they are processed via the amygdalar–hippocampal pathway. Nonarousing words processed via the prefrontal–hippocampal pathway should not produce more vivid memories than arousing words. No study has directly studied how manipulating levels of encoding influences recall of taboo words. The current study is designed to add to the existent body of knowledge about memory for taboo words by looking at encoding strategies and arousal levels used to process taboo words.

Taboo word recall rates should be less influenced by processing level than neutral words or valenced (positive or negative) words. Taboo word recall should be better than recall of nontaboo words, exhibiting the emotionality effect regardless of encoding strategy. Neutral words should

exhibit the LOP effect with better recall as a result of deep processing relative to shallow processing. Valenced words should show intermediate levels of benefit from LOP.

EXPERIMENT 1

METHOD

Participants

Participants were 18 undergraduate and graduate students (8 women, 10 men) at a large university in New York City. Their ages ranged from 19 to 39 years, with a mean age of 24 years. All participants volunteered to be in the study. Non–native English speakers were excluded from the study because past research has shown that taboo words elicit greater autonomic reactivity in a first rather than in a second language (Harris, Aycicegi, & Gleason, 2003).

Materials

Two types of orienting questions were used: shallow and deep. An example of a shallow question is, "Is the word in upper case?" An example of a deep question is, "Does this word fit in the sentence: *The* _____ *is blue?*" There were 36 stimulus words, which included 12 taboo words, 12 emotional words (6 positive, 6 negative), and 12 neutral words (Table 1). Words were chosen from MacKay et al. (2004), Harris et al. (2003), and Bellezza, Greenwald, and Banaji (1986). The taboo words are socially proscribed profanities, insults, and sexual references and received higher obscenity ratings than the neutral words (Jay, 1992). Stimulus words were matched for familiarity, number of syllables, and number of characters using MacKay et al.'s (2004) norms. The list of words and orienting questions was randomized.

Table 1. Recall as a function of word type for individual words, Experiment 1

Neutral word	Recall (%)	Emotional word	Recall (%)	Taboo word	Recall (%)
Wife	23.53	Friend	41.18	Nigger	64.71
Brother	17.65	Cuddle	29.41	Bitch	64.71
Host	11.76	Cozy	23.53	Pussy	58.82
Attic	5.88	Love	17.65	Cock	52.94
Bank	5.88	Fight	11.76	Slut	52.94
Cross	5.88	Sick	11.76	Anus	35.29
Note	5.88	Kill	5.88	Rape	35.29
Page	5.88	Hurt	5.88	Scrotum	29.41
Senate	5.88	Freedom	5.88	Piss	23.53
Frame	0.00	Anger	5.88	Dyke	23.53
Lung	0.00	Kiss	0.00	Queer	17.65
Pity	0.00	Crime	0.00	Shit	11.76

Procedure

Experimental trials consisted of a deep or shallow orienting question followed by the stimulus word. The question remained visible until participants indicated their comprehension by pressing the spacebar. Duration of the word targets was 170 ms, followed by a blank screen. Participants responded during this blank screen by pushing the "yes" or "no" key on the keyboard (letters "D" and "K"). Their response triggered appearance of the next question. Orienting questions and stimulus words appeared on a computer monitor in large, black font against a white background and were presented via the SuperLab experimental control software. The experiment began with participants reading instructions on the computer screen and completing six practice items.

After all 36 words were presented, participants were given a filled retention interval to stimulate forgetting in which an arithmetic worksheet had to be completed. After spending 10 min on the worksheet, participants received a surprise recall test. They were instructed to write down on a blank piece of paper as many words as they could recall from those seen during the prior computer task.

RESULTS

Participants recalled the taboo words significantly better (39%) than they recalled the emotional words (13%), $F(1, 17) = 67.12$, $p < .001$ (Figure 1). Participants recalled a marginally greater percentage of the emotional words than the neutral words (7%), $F(1, 17) = 3.89$, $p = .065$. These two results represent emotionality effects. Orienting task did interact with word type, $F(2, 34) = 3.70$, $p < .05$. Using a 2 (shallow, deep) × 2 (emotional, taboo) ANOVA, participants remembered a greater percentage of the neutral words with the deep task (10%) than the shallow task (5%), replicating the typical LOP effect. This was the opposite of what happened with the taboo words. Participants remembered significantly fewer with the deep task (34%) than with the shallow task (44%), $F(1, 17) = 7.25$, $p < .02$. Participant gender was not a significant variable, and gender did not interact with word recall, $F(1, 16) = 0.05$, *ns*. Errors of commission (i.e., reporting taboo words that were not on the stimulus list) were rare.

Table 1 shows the percentage of participants recalling individual words as a function of word type. Here one can see the superior recall of taboo words over other word types and also note the words within each category are not equally recallable. The taboo words that are recalled best are those that are highly taboo and are probably more arousing than less taboo words are. The correlation between recall rate and tabooness ratings from Jay (1992) is $r = .67$, $p < .02$.

Within the emotional word category, we separately analyzed positive and negative words to determine whether valence influenced recall (see Kuhlmann, Piel, & Wolf, 2005). The rates of recall for positive words were 19%

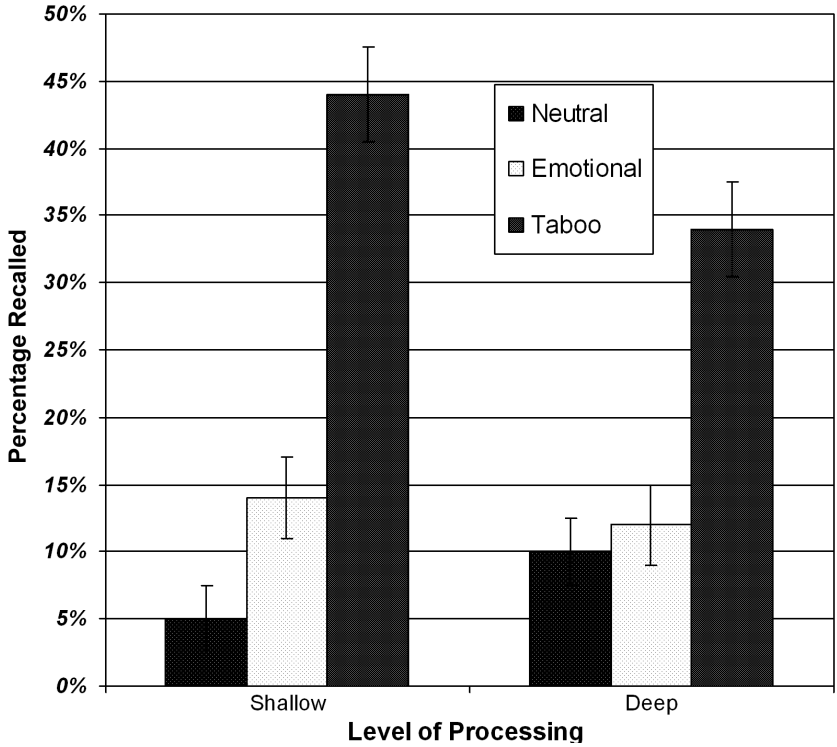

Figure 1. Percentages of word recall as a function of orienting task

and 14% (deep and shallow conditions, respectively), and for the negative words they were 6% and 14% (deep and shallow conditions, respectively), but these were not statistically significant differences, $F(1, 17) = 2.55$, $p > .15$. These positive and negative emotional words, as with emotional word stimuli used in previous memory research, are not as memorable as taboo words. Mean reaction times for taboo, positive, negative, and neutral words were 1,250 ms, 1,299 ms, 1,114 ms, and 1,180 ms, respectively, but these values were not statistically different, $p > .20$. Recall and response time also did not differ according to whether the correct response was a "yes" or a "no" response.

To investigate whether taboo words had priority during the recall phase, we analyzed the order in which words were recalled by assigning ranks to words recalled by each participant. The mean rank was highest for taboo words and lowest for neutral words. The mean rank for each category and its range, standard deviation, and frequency (across all participants) appear in Table 2. Histograms showed that taboo words had a large number of low-ranked values (ranks 1–4), with an exponentially decreasing

distribution (long rightward tail), with a small number of taboo words being recalled as the 13th or 14th item. In contrast, the rank order of recall for neutral words was fairly normally distributed. The distribution for emotional words also had a rightward tail but was intermediate between neutral and taboo words in terms of the normality of the distribution. Although the taboo words had far more early-ranked items than the other two categories, taboo words predominated over other items at every rank. The relative differences in recall as a function of word type suggest that participants did not suspect that they would have to recall the words from the computerized task; otherwise, recall levels would be more equivalent across word types. The recall order data thus show that taboo words were recalled early and often, that is, taboo words typically were the first words to be recalled, but taboo words predominated at every stage in recall. Consistent with this, during debriefing several participants made statements about the memorability of the taboo words. A typical comment was, "I had the feeling that I would be able to write down many words, but when I tried to remember them only curse words came to mind." No participants mentioned that they suspected a memory test would be used.

DISCUSSION

Results demonstrated evidence for both an LOP effect and an emotionality effect. An LOP effect was obtained for the neutral words but not for the taboo words or emotional words. Taboo word recall exhibited an emotionality effect. The lack of an LOP effect for emotional and taboo words is a striking effect that has not previously been reported in the literature.

Why would taboo and emotional words not benefit from deep processing? One explanation is that superior memories for taboo and emotional words are formed on the basis of their arousal levels and processing in the amygdalar–hippocampal pathway. But taboo and emotional words may also benefit more during retrieval. In Experiment 1 we did not manipulate participants' recall strategies. Experiment 2 was designed to examine encoding and retrieval factors involved in processing taboo and nontaboo stimuli. To determine whether emotional and taboo words form superior memories regardless of the shallow or deep task, we monitored partici-

Table 2. Recall by word category, Experiment 1

Category	Mean	Range	*SD*	Frequency
Neutral	6.5	1–15	3.9	15
Emotional	4.5	1–9	2.5	25
Taboo	4.6	1–15	3.0	83

pants' electrodermal activity. If words' emotionality produces arousal, then taboo and emotional words should elicit higher SCRs in both shallow and deep encoding conditions. We cannot rule out memory effects due to retrieval strategies, so to investigate retrieval factors, we varied recall instructions across participants. Some participants engaged in free recall (i.e., noncued), and some were asked to recall some categories of words before others.

EXPERIMENT 2

Experiment 1 demonstrated that the recall of taboo words and emotional words is based on emotionality and is not increased by deeper processing, contrary to the standard finding in the depth-of-processing literature. One explanation is that taboo words elicit sufficient emotional arousal that memory is enhanced regardless of encoding task (Anderson, 2005). To probe this explanation, we directly measured emotional arousal by monitoring SCRs during the initial study period. We predicted that SCRs to taboo words would be high, relative to emotional and neutral words, regardless of whether participants were performing a deep or shallow encoding task.

We additionally modified the LOP procedure in order to probe a curious aspect of the Experiment 1 results: Recall of taboo words was worse in the deep processing conditions (Figure 1). Could this have occurred because the deep processing task (a sentence congruity task) was one that drew attention away from taboo words, one that focused on the semantic features of the words rather than their emotionally arousing connotative properties? Consider what processing may occur when participants needed to judge the semantic congruity of the taboo word *pussy* as a completion for the sentence template, "*The* _____ *is blue.*" The sentence template detracts attention from the taboo, emotionally arousing connotations of *pussy* by placing it in a specific syntactic and semantic environment (i.e., focusing on *pussy* as a noun). Conversely, without an accompanying semantic task, as in the shallow case judgment task, arousing taboo associations may more easily come to mind, engaging amygdalar–hippocampal processing. For these reasons, we chose a different LOP task: the semantic categorization task used by Jacoby and Dallas (1981). In this task, participants verified that an indicated word was a member of a specific category (e.g., body part, emotion word). This form of the deep task should constrain the associations to the target words because the two orienting questions focus the readers' attention on a more limited range of semantics (body parts and emotions) than a procedure in Experiment 1 where many types of orienting questions were used (e.g., colors). That pool of questions shifted the semantic

frame on each trial. The shallow task used was the same as in Experiment 1 (to verify that a word had appeared in uppercase or lowercase).

We also sought to address whether the high recall for taboo words occurred because participants elected to recall these words first. Recall of taboo words would then diminish the ability to recall emotional and neutral words. Of course, the reason for recalling taboo words first is their emotional salience. The salience or arousal level of taboo words makes them more memorable at recall regardless of orienting task. However, one could imagine that participants recalled taboo words first for strategic reasons. To investigate this, we randomly assigned participants to a free recall condition (the method used in Experiment 1) or to one of three cued recall conditions. In each of these conditions, one of three categories (taboo words, emotional words, and animal words) was cued to be recalled first.

A final question was whether part of the recall advantage for taboo words occurs because they form a more coherent category than did the neutral word category, as recently argued by Talmi and Moscovitz (2004), who found an effect of semantic relatedness on recall in an incidental encoding paradigm (also see LaBar & Phelps, 1998; MacKay, Hadley, & Schwartz, 2005, on category cohesiveness). We followed these authors and used animal words instead of a semantically diverse neutral word category.

METHOD

Participants

Participants were 40 undergraduate students of Boston University (70% women, reflecting gender ratios in introductory psychology classes). Ages ranged from 18 to 22 years. As in Experiment 1, non–native English speakers were excluded because they have greater autonomic reactivity in a first than in a second language (Harris et al., 2003).

Materials

Participants were presented with two types of orienting questions: a shallow question and deep questions. The shallow question was the same as in Experiment 1 (i.e., "Is the word in uppercase?"). Twelve deep questions were used, with each question used two to four times across the 36 trials. Questions were constructed to activate semantic associates of the target word. For example, for animal targets, questions included "Is this a farm animal?" and "Is this a carnivore?" For taboo words, questions included "Is this a sexual act?" and "Is this a derogatory term?" For emotional words, questions included "Is this an affectionate act?" and "Is this an unpleasant sensation?"

The 36 stimulus words were composed of equal numbers of emotional words, taboo words, and animal words. The 12 animal words were chosen from MacKay et al. (2005), and the 12 taboo words were matched with nontaboo words for

length, number of syllables, and familiarity (using ratings provided by MacKay et al., 2004). The 12 emotional words were the same as the negative and positive terms used in Experiment 1. All stimuli appear in Table 3.

Procedure

The procedure and apparatus were similar to those in Experiment 1. Participants read instructions that appeared on the computer screen. The participants were instructed to first read a question that would appear on the screen. One of the two types of orienting questions was randomly assigned to each stimulus word. Instructed to keep the question in mind, the participants pressed the spacebar to receive the stimulus word. The word then appeared on the screen for 200 ms, followed by a blank screen for 7,000 ms. During the blank screen the participant responded by pushing the assigned "yes" or "no" key on the keyboard. After the 7-s interval the next question appeared on the screen. Six practice trials preceded the experimental trials.

Electrodermal monitoring and analysis

After the procedure was explained to them, participants were connected to skin conductance apparatus. A 7-s recording interval began coincident with stimulus onset. Gold-plated electrodes were attached to the tip of the index and middle fingers of the dominant hand. Electrodermal activity (tonic and phasic) was recorded using the Davicon C2A Custom Skin Conductance Monitor (NeuroDyne Medical Corporation). Neusoft software from NeuroDyne recorded skin conductance levels in micromhos. The amplitude of the phasic SCR was obtained by subtracting the base point, defined as lowest skin conductance level during the 7-s recording window, from the maximum score. Following Dawson, Schell, and Filion (2000), we calculated SCR frequencies for each condition. SCR frequency is the percentage of trials on which an SCR with an amplitude of 0.05 micromhos

Table 3. Recall as a function of word type for individual words, Experiment 2

Animal word	Recall (%)	Emotional word	Recall (%)	Taboo word	Recall (%)
Cow	46.2	Love	51.2	Fuck	77.5
Turtle	42.5	Anger	43.9	Nigger	61.0
Mink	40.0	Cuddle	34.1	Pussy	60.5
Deer	39.0	Kill	25.0	Ass	58.5
Skunk	37.5	Cozy	25.0	Slut	55.8
Sheep	35.0	Kiss	23.1	Chink	55.0
Mouse	34.2	Hurt	19.5	Whore	48.6
Rat	34.1	Crime	15.4	Tit	43.6
Horse	33.3	Friend	15.0	Screw	38.5
Panther	32.4	Freedom	12.5	Rape	37.5
Bear	25.0	Fight	12.5	Bitch	32.6
Tiger	22.5	Sick	5.1	Bastard	7.5

occurs. That is, we assigned trials a 1 if the SCR was greater than 0.05 micromhos or a 0 if SCR amplitude was less than 0.05 (Eisenberg et al., 1991).

After all 36 questions and stimulus words were presented, participants were given a filled retention interval that used an intervening task to stimulate forgetting. The intervening task was to complete the Behavioral Inhibition System/Behavioral Activation System scale. This is a self-report questionnaire sensitive to the personality constructs of inhibition. We used it to allow us to factor individual variation in emotional arousal into our analysis of SCRs. Participants were then given a surprise recall task and were randomly assigned to one of the following four recall conditions:

Free recall. The written instructions on the sheet of paper asked participants to write down all the words recalled from the prior computer task.
Animal first, then emotional, then taboo.
Taboo first, then animal, then emotional.
Emotional first, then taboo, then animal.

Combining over these last three tasks, each category has an opportunity to be recalled first, second, and third. We will use the "animal first" condition to illustrate the written instructions. The instructions noted that the prior task had contained three categories of words: animal words, emotional words, and taboo words. Participants were instructed to write down all the animal words they recalled. When they had finished, the experimenter turned over the piece of paper, and written instructions directed them to write down as many emotional words as they could recall. When they were finished with this, the experimenter handed them a new sheet of paper, which directed them to write down as many taboo words as possible.

RESULTS

Two dependent measures were of interest: percentage recall and skin conductance amplitudes.

Percentage recall of words

Averaging over the four recall conditions, a 2×3 ANOVA, with within-participant factors depth of processing (shallow vs. deep) and word category (taboo, emotional, and animal word), revealed main effects of depth of processing and word category. More items were recalled with deep than shallow processing, $F(1, 39) = 155$, $p < .0001$. The main effect of word category was significant at $F(2, 78) = 31.5$, $p < .01$. Confirming the findings in Experiment 1, recall rates were higher for taboo words (mean 48%) than for animal words (35%) and emotional words (24%). The interaction was also significant, $F(2, 78) = 4.3$, $p < .02$. Post hoc Bonferroni *t*-tests comparing deep and shallow recall performance were significant for all word categories: animal, $t = 10.65$; taboo, $t = 5.16$; negative emotional, $t = 5.43$; and positive emotional, $t = 4.61$. As shown in Figure 2, animal words (our neutral category) benefited the most from deep processing

(53% – 17% = 36%). In contrast, the difference between deep and shallow processing was 23% for emotional words and 22% for taboo words. In an exploratory analysis, we categorized emotional words as having positive or negative valence. There was a nonsignificant trend for positive words to have higher rates of recall than negative words, $p < .08$. We quantified the advantage of deep processing for each word category by calculating effect sizes (partial η^2, obtained via SPSS repeated-measures ANOVA). These were .76, .48, .43, and .37 for animal words, negative words, taboo words, and positive words, respectively. These effect sizes show that the animal words benefited more from deep encoding than the emotional and taboo words.

We examined how recall cues influenced recall with a three-way ANOVA (recall order × word category × level of processing). The three-way interaction was not significant, $F < 1$, although a highly significant interaction of recall order and word category was found, $F(6, 72) = 5.5$, $p < .001$. This interaction is graphed in Figure 3. Condition labels indicate the order of recall cueing (e.g., "taboo–animal" indicates that participants were first instructed to recall taboo words, then animal words, and then emotional words). Recall performance varied as a function of word category for all recall conditions except when emotional words were cued to be recalled first. Significant differences were found for free recall, $F(2, 16) = 23.70$; taboo first, $F(2, 16) = 24.30$; and animal first, $F(2, 16) = 10.6$; but not for the emotional first condition, $F < 1$, *ns*. Post hoc Scheffe tests revealed that overall recall was greater in the taboo first condition (62%) than in

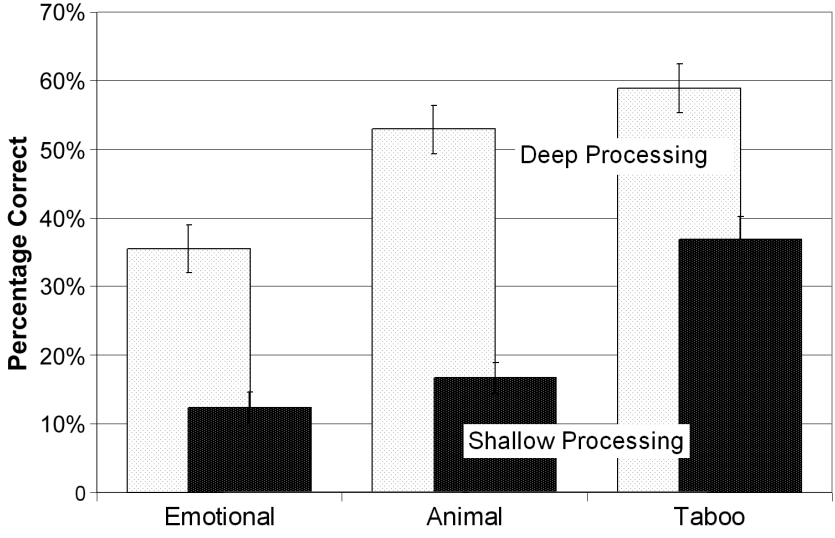

Figure 2. Percentage correct, Experiment 2

the animal first condition (36%), $p < .03$. No other pairwise comparisons were statistically different. Recall of taboo words was significantly greater in the taboo first condition (62%) than in the emotional first condition (41%), $p < .05$ using a Scheffe test. Planned comparisons were conducted to determine whether taboo words were recalled better than the other word categories in each recall condition. Taboo words were recalled higher than the other categories in the taboo first and free recall conditions, all Fs$(1, 9) > 15$, ps $< .05$, but not in the emotional first and animal first categories.

The recall data for individual words are presented in Table 3. As in Experiment 1, the recall of taboo words was related to level of tabooness. The correlation between recall rates for taboo words and their tabooness ratings from Jay (1992) was $r = .63$, $p < .03$.

Electrodermal measures

A 2 × 3 ANOVA with factors level of processing (shallow vs. deep) and word category (taboo, emotional, and animal word) was conducted with SCR frequencies (meaning presence or absence of the SCR during the 7-s trial) as the dependent measures. Frequency of SCRs varied by word category, $F(1, 117) = 3.6$, $p < .02$, but no interaction occurred with task. As shown in Figure 4, taboo items elicited a higher frequency of SCRs in

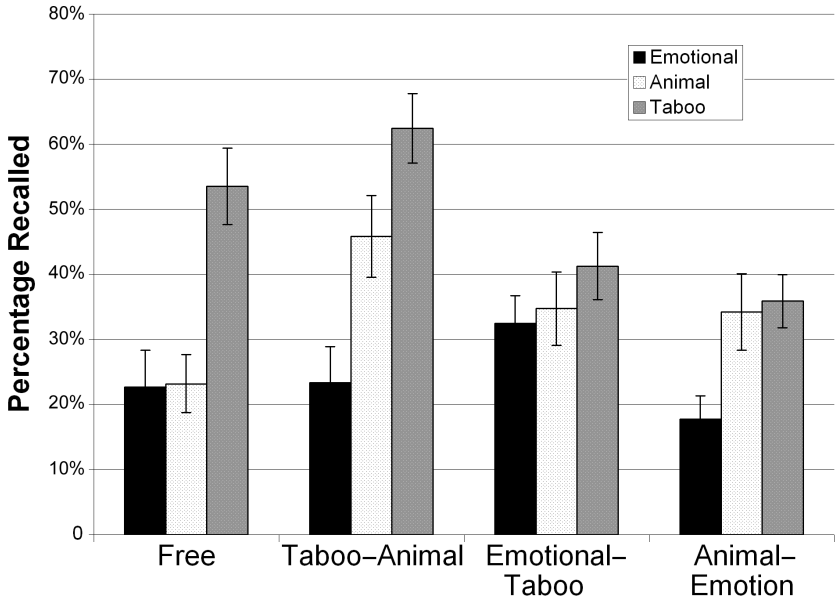

Figure 3. Percentage recalled for three word categories (emotional, animal, and taboo) as a function of the four recall conditions

both shallow and deep tasks than did emotional and animal words. Greater arousal thus matched percentage recall. Frequency of SCRs in each of the four categories (taboo, positive, negative, animal) correlated with percentage recall at $r = .85$, $p < .01$. The high correlation between SCR and overall recall may replicate the findings of Kleinsmith and Kaplan (1963), who found a memory advantage for nonsense syllables associated with SCRs.

DISCUSSION

This experiment replicated two key findings from Experiment 1: Taboo words had the highest recall, and the emotionally neutral category benefited more from deep processing than did the emotional and taboo words. Our results also confirmed that taboo words more frequently elicit

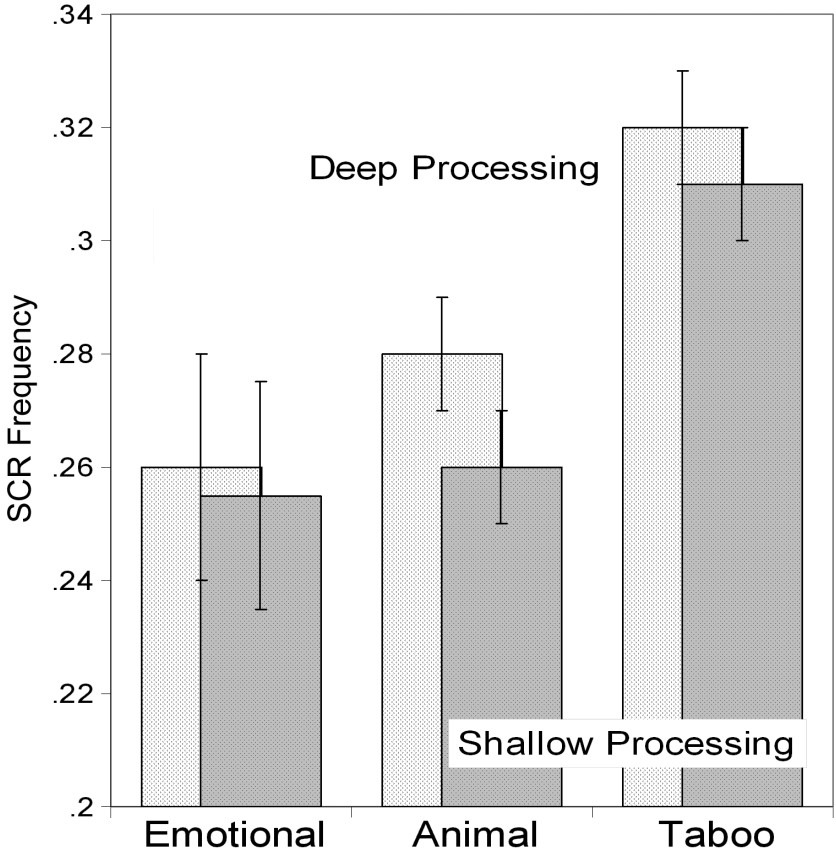

Figure 4. Skin conductance response (SCR), Experiment 2

skin conductance responses than the other words, regardless of processing level. Our results are consistent with those of Sharot and Phelps (2004), indicating that after a delay interval, arousing taboo words are remembered better than nonarousing nontaboo words. Other aspects of the experimental design were novel and exploratory.

Shallow and deep instructions

We wondered whether using a different deep encoding task from Experiment 1, category verification, would reveal the standard memory advantage for deep encoding (LOP effect) that was absent for taboo words and emotional words in Experiment 1. Taboo and emotional words did show an LOP effect, confirming that recall is enhanced with deeper elaboration even for the highly arousing taboo words. The important point is that the size of the LOP effect was smaller for taboo words (22%) than for the neutral category (36%).

Animal words as the neutral category

Averaged across conditions, animal words had higher recall than emotional words (35% vs. 24%). There are probably several reasons for this. We had selected animal words as our nonemotional category, given that previous researchers had identified animal terms as forming a highly cohesive category (e.g., MacKay et al., 2004; LaBar & Phelps, 1998). Animal words share many semantic features and occur thematically together in children's stories and movies. We suggest that the high category cohesiveness of animal words is one reason for the good recall. A second reason may be that animal words have higher concreteness and imageability than do emotional words (see Toglia & Battig, 1978). Specific evidence of the role of cohesiveness in increasing recall is the generally lower recall of animal words in the free recall condition (25%) than in the other recall conditions. In the free recall condition, participants were not informed that words in the study list could be grouped into specific categories. Recall of animal words in the free recall condition was 25%, which was lower than animal recall in the taboo first condition (45%) and significantly lower than the animal first condition (36%) but nonsignificantly lower than in the emotional first condition (34%). A third reason for the superior recall of animal words than emotional words concerns the context in which they are presented. In a study list containing taboo words, animal words may be interpreted as having taboo qualities, given that animal words can be considered insults (e.g., cow, skunk, rat).

Cued vs. free recall

One reason for the superior recall of taboo words might be that they came to mind first during free recall, as revealed by the order-of-report

analysis in Experiment 1. When taboo words were cued to be recalled first, the highest overall rates of recall were obtained (44%). This suggests that part of the reason for taboo words' superior recall is that they have formed vivid memories and they spontaneously come to mind first during the recall task. Inspection of Figure 3 suggests that the free recall condition was most similar in its pattern to the taboo first condition. This suggests that the free recall may function as a taboo first condition. In the other cued recall conditions, participants must avoid answering taboo words in order to first recall animal or emotional words. This strategy could invoke inhibitory mechanisms, following Gernsbacher's findings that instructions to recall specific items result in attempts to suppress recall of competing items (Gernsbacher & Faust, 1991). The expected result is that cueing nontaboo words will inhibit recall of taboo words, thus reducing their recall and reducing the total number of words recalled. This is what occurred. Total recall was lowest with animal words recalled first (30%) and highest with taboo words recalled first (43%). In contrast, the emotional first condition produced little advantage for emotional words, suggesting again that they did not form a cohesive category. In the context of vivid animal words and offensive taboo words, the emotional words apparently were less salient and less arousing. In stimulus sets without arousing taboo words, the emotional words could well be more memorable than they are here. The cohesiveness of taboo words as a category has not been established by psycholinguistic research.

GENERAL DISCUSSION

It has long been recognized that emotional words have superior recall to neutral words (Rapaport, 1942). However, until recently there were few explanations for why this occurs. These experiments tested the following explanation: The emotional connotations of taboo and emotional words are arousing, so that memory for taboo words is good regardless of task instructions. A specific hypothesis is that emotional words will benefit less when research participants are given tasks that induce deeper processing (Bower, 1981). This hypothesis was confirmed. In Experiment 1, taboo and emotional words did not show a recall advantage when deep encoding instructions were provided, although neutral words did show the classic deep encoding advantage. Asking participants to focus on the denotative aspects of taboo words was not beneficial. In Experiment 1, we found that one specific deep task, the sentence verification task, actually reduced recall for taboo and emotional words, compared with the shallow task (uppercase vs. lowercase identification). This might have occurred because the sentence verification task draws attention from the emotional

connotations of the words that are arousing by forcing participants to focus on less arousing denotative aspects of their meanings. When a different deep task was used, the category verification task (Experiment 2), an LOP effect for emotional and taboo words was found. In Experiment 2, taboo and emotional words benefited less from deep encoding than did neutral words. These findings indicate that connotations of emotional and taboo words are arousing and memorable even under shallow processing instructions. Whereas Bower's theory (1981) acknowledged the power of emotion to effect the memorability of words, the LOP approach does not directly address the role of emotion in word encoding.

The current research contributes to the LOP paradigm by demonstrating how a semantic encoding task does not invariably improve recall over a shallow task. For words that are arousing during word reading, deep encoding instructions are not necessary for facilitating memory. Furthermore, the nature of the task influences the strength of the emotionality advantage.

Skin conductance was monitored in Experiment 2 to test the hypothesis that taboo words are sufficiently inherently arousing that they do not benefit from deep encoding instructions. This hypothesis was confirmed: Taboo words had heightened frequency of SCRs in both the shallow and the deep encoding tasks.

It was not anticipated that animal words would be better recalled than emotional words. Using animal words as our neutral stimuli served the purpose of having a category that could be easily cued for recall. Additionally, this allowed us to have stimuli that formed a coherent category, just as the taboo words form a coherent category based on their negative valence and arousal level. Emotional words and taboo words show a lower LOP effect. We concluded that these words' inherent arousal levels elicit better memories independent of task instructions.

Notes

Correspondence about this article should be addressed to Timothy Jay, Department of Psychology, Massachusetts College of Liberal Arts, 375 Church Street, North Adams, MA 01247 (e-mail: t.jay@mcla.edu).

References

Adolphs, R., Russell, J., & Tranel, D. (1999). A role for the human amygdala in recognizing emotional arousal from unpleasant stimuli. *Psychological Science, 10,* 167–171.

Anderson, A. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology: General, 134,* 258–281.

Anderson, A., & Phelps, E. (2001a). Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. *Journal of Cognitive Neuroscience, 14,* 709–720.

Anderson, A., & Phelps, E. (2001b). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature, 411,* 305–309.

Bellezza, F. S., Greenwald, A. G., & Banaji, M. R. (1986). Words high and low in pleasantness as rated by male and female college students. *Behavior Research Methods, Instruments, & Computers, 18,* 299–303.

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36,* 129–148.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5,* 73–99.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Dawson, M., Schell, A., & Filion, D. (2000). The electrodermal system. In J. Cacioppo, L. Tassinary, & G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 200–223). New York: Cambridge University Press.

Eisenberg, N., Fabes, R. A., Schaller, M., Miller, P., Carlo, G., Poulin, R., et al. (1991). Personality and socialization correlates of vicarious emotional responding. *Journal of Personality and Social Psychology, 61,* 459–470.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 245–262.

Grady, C. L., McIntosh, A. R., Rajah, M. N., & Craik, F. I. M. (1998). Neural correlates of the episodic encoding of pictures and words. *Proceedings of the National Academy of Science, USA, 95,* 2703–2708.

Harris, C. L., Aycicegi, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first than in a second language. *Applied Psycholinguistics, 4,* 561–578.

Isenberg, N., Silbersweig, D., Engelien, A., Emmerich, S., Malavade, K., Beattie, B., et al. (1999). Linguistic threat activates the human amygdala. *Proceedings of the National Academy of Science, 96,* 10456–10459.

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

Jay, T. (1992). *Cursing in America.* Philadelphia: John Benjamins.

Jay, T. (2000). *Why we curse: A neuropsychosocial theory of speech.* Philadelphia: John Benjamins.

Kapur, S., Craik, F. I. M., Tulving, E., Wilson, A., Houle, S., & Brown, G. (1994). Neuroanatomical correlates of encoding in episodic memory: Levels of processing effect. *Proceedings of the National Academy of Science, 91,* 2008–2011.

Kensinger, E., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition, 31,* 1169–1180.

Kensinger, E., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural process for valence and arousal. *Proceedings of the National Academy of Sciences, 101,* 3310–3315.

Kern, R., Libkuman, T., & Otani, H. (2002). Memory for negatively arousing and neutral pictorial stimuli using a repeated testing paradigm. *Cognition & Emotion, 16,* 749–767.

Kern, R., Libkuman, T., Otani, H., & Holmes, K. (2005). Emotional stimuli, divided attention, and memory. *Emotion, 5,* 408–417.

Kleinsmith, L. J., & Kaplan, S. (1963). Paired associate learning as a function of arousal and interpolated interval. *Journal of Experimental Psychology, 65,* 190–193.

Kuhlmann, S., Piel, M., & Wolf, O. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *Journal of Neuroscience, 25,* 2977–2982.

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

MacKay, D. G., & Ahmetzanov, M. V. (2005). Emotion, memory, and attention in the taboo Stroop paradigm: An experimental analog of flashbulb memories. *Psychological Science, 16,* 25–32.

MacKay, D. G., Hadley, C. B., & Schwartz, J. H. (2005). Effects of emotion on orthographic repetition blindness and the illusory word phenomenon: Tests of binding theory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 58,* 1514–1533.

MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D., Abrams, L., & Dyer, J. (2004). Relations between emotion, memory and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition, 32,* 474–488.

Manning, S. K., & Melchiori, M. P. (1974). Words that upset urban college students: Measures with GSRs and rating scales. *Journal of Social Psychology, 94,* 305–306.

Maratos, E., Allan, K., & Rugg, M. (2000). Recognition memory for emotionally negative words: An ERP study. *Neuropsychologia, 38,* 1452–1465.

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

Rapaport, D. (1942). *Emotions and memory.* New York: International Universities Press.

Reber, R., Perrig, W., Flammer, A., & Walter, D. (1994). Levels of processing and memory for emotional words. *Schweizerische Zeitschrift für Psychologie, 53,* 78–85.

Sharot, T., Delgado, M., & Phelps, E. (2004). How emotion enhances the feeling of remembering. *Nature Neuroscience, 7,* 1376–1380.

Sharot, T., & Phelps, E. (2004). How arousal modulates memory: Disentangling the effects of attention and retention. *Cognitive, Affective, & Behavioral Neuroscience, 4,* 294–306.

Talmi, D., & Moscovitz, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory & Cognition, 32,* 742–751.

Toglia, M., & Battig, W. F. (1978). *Handbook of semantic word norms.* Hillsdale, NJ: Erlbaum.

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