

Effects of Prior Experience on Judgments of Normative Word Frequency: Automatic Bias and Correction

Jeffrey P. Toth and Karen A. Daniels

Georgia Institute of Technology

Four experiments examined repetition priming effects on judgments of normative word frequency. Experiment 1a found reliable priming, but only when words were studied under conditions of full attention; divided attention eliminated the effect. Experiment 1b replicated the attention effects and found that frequency and recognition judgments could be dissociated as a function of study repetition. Experiment 2 showed that Levels of Processing could also dissociate the two judgments. Posthoc latency analyses of the data from Experiments 1 and 2 showed the biasing of frequency judgments to be limited to the subjects' fastest responses; slower responses showed no bias effects. Experiment 3 replicated the latency effects using an instructional manipulation of response speed. Posttest interviews in all of the experiments suggested that the bias correction seen in the slower responses occurred without awareness or intent. The results extend prior research on repetition priming to a semantic judgment about a well-known class of stimuli and are discussed in terms of processing fluency, source monitoring, and the correction of unconsciously influenced judgments. © 2002 Elsevier Science (USA)

Key Words: implicit memory; word-frequency judgments; judgment bias; bias correction.

Past experience can affect subsequent performance in a variety of different ways. In addition to traditional expressions of memory such as recall and recognition, people also show effects of prior experience that are unaccompanied by conscious recollection of the past, and these effects may manifest themselves as changes in perception, categorization, reasoning, or motor behavior (Bargh, 1997; Kelley & Lindsay, 1996; Roediger & McDermott, 1993; Toth, 2000). Much of the research examining such "implicit" effects has been directed at showing how they differ from more conscious or "explicit" expressions of memory. This research agenda—demonstrating dissociations between implicit and explicit tests—has been important in establishing the existence of two, if not more, different kinds of memory. However, a drawback of this agenda is that it fails to capture one of the more important facts about memory, namely, that these two or more forms of

memory coexist in the mind of the rememberer. That is, treating conscious and unconscious forms of memory as isolated phenomena, as one is compelled to do when employing implicit and explicit tests, ignores the interplay between them and thus the experiential qualities and judgment strategies that emerge when these two forms of memory are operating together.

While not denying the qualitative distinction between conscious (explicit) and unconscious (implicit) forms of memory, we have become interested in how these two forms of memory interact to create different subjective experiences and thus different influences on performance. The experiments presented here provide some insight into the nature of this interaction. The experiments began as an attempt to produce unconscious influences of memory in a judgment domain for which people have had a great deal of prior experience. Subjects were first exposed to a list of common English words and then were asked to judge the normative (background) frequency of these words, as well as others that had not been previously presented. Instructions for the judgment task were similar to those used in other implicit tests; that is, subjects were led to believe that their judgments were unrelated to the previous encoding experience. As will be seen, the results provide relatively clear evi-

We thank Colleen Parks and JoEllen Cox for their help in data collection. We also thank Reed Hunt, Steve Lindsay, Larry Jacoby, Colleen Parks, Randy Engle, and an anonymous reviewer for their insightful comments on a previous version of this paper.

Address correspondence and reprint requests to Jeffrey P. Toth, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332-0170. E-mail: jt104@prism.gatech.edu.

dence that such implicit influences are indeed possible—that a single prior experience can bias the apparent normative frequency of English words. However, the results also suggest that subjects were sensitive, at some level, to the nature of this bias and, under some conditions, were able to correct for it. As will be discussed in more detail below, we believe this interplay between bias and bias correction may be a fundamental characteristic of a variety of real-world judgment tasks that are susceptible to unconscious influences of memory.

Prior to describing the experiments, we first provide more background information on our theoretical approach to implicit memory and then explain why judgments of normative word frequency present a particularly interesting test of this form of memory.

UNCONSCIOUS INFLUENCES AND THE FLUENCY HEURISTIC

Along with other researchers (e.g., Jacoby & Brooks, 1984; Jacoby & Dallas, 1981; Kolers & Roediger, 1984; Masson, 1989; Whittlesea, 1997), we view implicit forms of memory as instances of fluent reprocessing. That is, encoding of an experience establishes a pathway for perceiving and/or conceptualizing such that subsequent encounters with repeated aspects of that prior experience result in processing that is faster and more efficient. This increased fluency of processing can have multiple effects on thought and behavior.

One of the most interesting aspects of fluent processing concerns its effects on subjective experience and the judgments based on such experiences (see Jacoby, Kelley, & Dywan, 1989b). For example, research has shown that fluent processing as a function of prior experience can increase the apparent fame of nonfamous names (Jacoby, Kelley, Brown, & Jasechko, 1989a), can lengthen the apparent exposure duration of briefly flashed words (Witherspoon & Allan, 1985), can lower the apparent loudness of background noise (Jacoby, Allan, Collins, & Larwill, 1988), and can increase the apparent truth of false statements (Begg, Anas, & Farinacci, 1992). These results suggest that fluent processing is a rather nonspecific (“sourceless”) form of information that can be transformed into a vari-

ety of different subjective experiences. The critical factor in this transformation appears to be the context in which fluent processing occurs, including the question/judgment asked of the subject (see Whittlesea, 1993; for an important elaboration of the fluency hypothesis, see Whittlesea & Williams, 1998, 2000).

The present experiments addressed two inter-related issues about the nature of processing fluency and its effects on judgment. The first concerns the kind of materials and judgment tasks that lend themselves to memory-based distortions and is motivated by the observation that much of the prior research has employed relatively novel stimuli and/or judgment tasks. For example, in two of the most popular judgment tasks—the false-fame and false-truth paradigms—the stimuli employed are ones to which subjects have not likely been previously exposed (e.g., “Sebastian Weisdorf”; “The temperature on the moon falls to 116 degrees below zero at night”). Thus an important question, for both theoretical and applied purposes, is whether memory-based distortions of judgments occur for stimuli with which people have had a great deal of prior experience. Given the research cited above, one might expect that judgments of well-known stimuli would be as susceptible to fluency-based influences as novel stimuli. But consider the task of judging the normative frequency of common English words such as “context,” “factory,” and “illness.” All of these words are highly familiar to native English speakers; indeed, it is likely that such speakers have had hundreds, if not thousands, of exposures to these words in their lifetime and also likely that they have been exposed to these words in the relatively recent past. Thus, if a person was able to restrict their judgment to semantic memory (should such a memory system exist), then it seems unlikely that they would be influenced by a single prior presentation. Moreover, if it could be shown that judgments about well-known stimuli are immune to implicit bias effects, it would limit the relevance of such effects to everyday behavior.

The second issue we address about fluency-based distortions concerns the different strategies that subjects have available to make their judgments. A notable consequence of experi-

ments employing novel stimuli is that there are few alternative sources a subject can draw on to make their judgment, other than the fluency of their processing. For example, without the aid of oscilloscopes or sound-meters, subjects in duration (Witherspoon & Allan, 1985)- and loudness-judgment experiments (Jacoby et al., 1988) have only their subjective perceptual experience on which to base their responses, thus making them susceptible to subtle differences in processing fluency (see Kelley & Jacoby, 1996, for related discussion). Similarly with the false-fame effect; although recollection that a name was presented on a "nonfamous" study list can allow subjects to discount the distorting (fame producing) effects of familiarity (e.g., Malthaup, 1995), in the absence of such recollection subjects have no alternative bases on which to make their judgment. In this regard, judgments of normative word frequency are interesting because, given subjects' vast exposure to word stimuli, they could potentially employ a variety of more analytic strategies to make their judgments (a point we will expand upon below).

Thus, the use of normative word-frequency judgments is an important extension of implicit memory and the fluency hypothesis, for three reasons. First, it extends implicit memory and fluency-based phenomena to a previously unexplored form of judgment. Second, the word-frequency task represents a relatively clear case in which, in addition to processing fluency, subjects have alternative, analytic bases for their judgments. Finally, to the extent that implicit biases are found, the results would have direct implications for the use of frequency judgments in applied contexts, as well as for more general theories of frequency estimation.

We should note that fluent reprocessing is not the only mechanism that could explain repetition priming effects on judgments of normative frequency. In particular, such effects could also potentially be explained as a form of response bias (Ratcliff & McKoon, 1996) or, as explored more fully in the General Discussion, as a failure of source monitoring (Johnson, Hashtroudi, & Lindsay, 1993). We emphasize the fluency heuristic here because of its role in motivating us to perform the present experiments and because of the similarity between frequency judg-

ments and other evaluative judgments (e.g., fame, truth, preference, pleasantness) for which processing fluency has been used as an explanatory mechanism (e.g., Begg et al., 1992; Jacoby et al., 1989a; Seamon, McKenna, & Binder, 1998; Whittlesea, 1993).

PRIOR RESEARCH ON FREQUENCY JUDGMENTS

Frequency judgments of prior word presentations, as well as other events, have a long history in psychology and have been important in both theoretical and applied contexts. Theoretically, such judgments have been used to address a number of questions, including how separate but related events are stored in memory (i.e., as a single, strength-sensitive representation or as separate instances: see Hintzman, 1988, 2000; Hintzman & Block, 1971; Howell, 1973), whether frequency of occurrence is represented as a distinct attribute (e.g., Begg, Maxwell, Mitterer, & Harris, 1986; Jonides & Jones, 1992), and whether this attribute (or some other memorial factor supporting frequency judgments) is encoded automatically (Hasher & Zacks, 1979, 1984). As well, frequency judgments have been theoretically related to judgments of recency (e.g., Huppert & Percy, 1978) and to the temporal dating of real-world events (Brown, Rips, & Shevell, 1985). Frequency judgments have also been investigated in applied contexts, for example, in consumer and survey research, where subjects are often asked to estimate the frequency with which they have previously performed various behaviors (e.g., Burton & Blair, 1991; Means & Loftus, 1991; Menon, 1993). A general conclusion from all of this research is that subjects are surprisingly accurate at judging the frequency of incidentally encoded events (see Hasher & Zacks, 1984; Hintzman, 2000; Howell, 1973), although a number of systematic biases are also apparent (e.g., Lichtenstein, Slovik, Fischhoff, Layman, & Combs, 1978; Tversky & Kahneman, 1973).

Although there is not universal agreement on how frequency judgments are rendered, prior research suggests that a variety of strategies could be employed depending on specific task demands (e.g., the type and periodicity of the event being evaluated, how the to-be-estimated events are

represented in memory, etc.). Consistent with this multiple-strategy perspective, Brown (1995) categorized frequency-estimation strategies into two general classes, enumeration and nonenumeration. The former involve explicit retrieval of prior relevant instances or events, which are then explicitly counted. The count achieved is then output or used as the basis for an extrapolated estimate. In contrast, nonenumeration strategies produce a frequency estimate through either (a) the explicit retrieval of a frequency count (presumably automatically computed during event presentation: Hasher & Zacks, 1979) or (b) an indirect, heuristic assessment of the contents of memory, based on characteristics such as availability, strength, or similarity to other instances in memory. Given the high-exposure rate for the words used in our experiments, a strict enumeration strategy (explicit retrieval and counting) seems unlikely, as does the retrieval of an automatic counter. Rather, we assume that subjects predominantly used an indirect, heuristic approach to our task. Indeed, evidence for the use of such a "indirect" strategy would seem to be provided by the existence of single-exposure, repetition priming effects on judgments of normative frequency, effects that we consistently demonstrate across the experiments reported here.

The current experiments differ from prior research on frequency judgments in a variety of ways. For example, most earlier studies of word-frequency judgments have been concerned with *situational* word frequency—that is, judgments of how often a word has been presented in a specific (e.g., experimental) context (for reviews, see Howell, 1973; Hasher & Zacks, 1984; Hintzman, 2000; Hock & Hasher, 1990). Importantly, judgments of situational word frequency are clearly "explicit" in the sense that subjects are made fully aware that words were presented different numbers of times in a previous encoding task and are asked to intentionally remember the frequency of those prior presentations. We break with this earlier research by examining how prior presentation may have unconscious/implicit influences on judgments of *normative* ("background") word frequency.

Also relevant to, but distinct from, the present research are studies examining the judged frequency of real world issues and events. For ex-

ample, Lichtenstein et al. (1978) had subjects estimate the frequency of lethal events (e.g., homicide, diabetes, etc.). Similarly, Brown and Seigler (1992) examined subjects' ability to estimate the size of national populations. These studies are similar to the present research in that both required judgments about a "semantic" class of knowledge. However, relative to judgments of normative word frequency, these earlier experiments differ critically in terms of the amount and type of prior experience that subjects could bring to bear on the various judgments. In particular, it seems likely that subjects in the Brown and Seigler (1992) and Lichtenstein et al. (1978) studies had little or no *direct experience* with the category being judged; that is, no subjects had direct exposure to the people making up, for example, the population of Bulgaria or to the number of people who have died from poisoning. In contrast, all English speaking subjects have had direct experience with the to-be-judged words used in the present experiments. We believe this difference is important, especially if one's interest is in how unconscious influences of memory could act to bias everyday judgments about well-known stimuli. That is, extensive, direct experience with a class of stimuli can be reasonably expected to produce a richer, more organized knowledge base that could potentially be used to reduce or eliminate extraneous influences on judgment (Srull, 1983).

To summarize, implicit memory, perhaps operating via a fluency heuristic, has been shown to distort a variety of evaluative judgments. An important question concerns the range of stimuli and judgments for which such fluency-based distortions may occur. The present experiments examined this issue in the context of a normative word-frequency judgment task. Our interest was not in the exact mechanisms underlying frequency judgments per se. Rather, we chose judgments of normative word frequency because of subjects' extensive prior experience with the to-be-judged stimuli, experience which could, in principle, be used to avoid the distorting effects of implicit memory. Such judgments thus constitute an interesting test case for implicit memory because of evidence showing that subjects can apply multiple strategies in making frequency estimations and judgments (see Brown, 1995).

IMPLICIT MEMORY AND INTENTIONALITY

A final issue to be addressed concerns our ability to clearly identify implicit influences as the basis for any obtained priming effects. Although the absence of conscious, intentional recollection is considered to be the defining characteristic of implicit memory (Schacter, 1987), unequivocally establishing the lack of conscious intent is by no means a straightforward matter (see Richardson-Klavehn & Bjork, 1988; Roediger & McDermott, 1993; Reingold & Merikle, 1990). Thus, although the present experiments employed implicit test instructions, as well as posttest interviews designed to detect explicit uses of memory, it is unclear whether such procedures can ensure that subjects only used automatic, implicit uses of memory at the time of judgment. As noted by a number of researchers (e.g., Bowers & Schacter, 1990; Jacoby, 1991; Toth, Reingold, & Jacoby, 1994), implicit instructions preclude neither spontaneous awareness of the prior encoding experience nor intentional uses of memory. Similarly, it is unclear whether retrospective reports can validly identify aware or intentional response strategies on a previous test, because they depend on the accuracy of memory for prior (and perhaps fleeting) states of awareness and intent. As well, there is extensive evidence that the nature of interview questions themselves influence subjective reports of prior states, dispositions, and behaviors (Schwarz, 1999).

Given the above issues, we employed two additional techniques to provide supporting evidence for the operation of implicit influences. First, we attempted to conform as closely as possible to tenets of the retrieval intentionality criterion (Schacter, Bowers, & Booker, 1989). This criterion requires that all aspects of an implicit and explicit test be equated (i.e., type of retrieval cue and response options) with only the retrieval instructions varied. If the same experimental manipulations are applied to both tests, then any differences in performance can be attributed to different retrieval processes. In the current study, we manipulated attention and study repetitions at encoding in Experiments 1a and 1b, study elaboration (level of processing) in Experiment 2, and study repetition and re-

sponse speed in Experiment 3. More importantly, Experiments 1b, 2, and 3 examined recognition memory, in addition to frequency judgments, and the stimuli presented for these two judgments were identical (i.e., single words that stayed available until response), as was the type of response (i.e., a binary choice). All that differed between the two tests was the goal of the judgment (normative frequency or recognition memory) and, more importantly, whether the instructions requested participants to use (recognition test) or not use (frequency test) episodic memory for a word's prior presentation. Thus, Experiments 1b–3 conformed closely to the retrieval intentionality criterion. To the extent that frequency and recognition judgments are found to be differentially affected by experimental variables, we would have evidence for the claim that subjects were not simply treating the frequency-judgment task as a test of episodic memory and, conversely, that judgments of normative frequency can be influenced by unintentional (implicit) uses of memory.¹

The second technique used to assess the implicit nature of any frequency bias was to examine performance as a function of response la-

¹As noted by a reviewer, given the difference between what subjects are asked to judge in the recognition and frequency tasks (old vs. new in the former, high vs. low in the latter), it is debatable whether our contrast of the two tasks constitutes an appropriate application of the retrieval intentionality criterion (RIC). We are sensitive to this criticism, but believe that the contrast between frequency and recognition judgments is interesting and relevant in the present experiments for two reasons. First, one of the main goals of the RIC is simply to establish that subjects are not employing explicit retrieval strategies on an implicit test (or, at least, are employing such strategies to a lesser extent). Given that our frequency and recognition tests employed identical retrieval cues, we believe that dissociations between the two tasks provide converging evidence for this conclusion (i.e., that subjects did not treat the frequency task as one of explicit recognition). A second, stronger argument for the appropriateness of the RIC is based on the view that recognition judgments are, in fact, frequency judgments—that is, judgments of whether a test word has a situational/study-list frequency of 1 or higher (see Hintzman, 1988; Hintzman & Curran, 1994). From this perspective, then, the frequency/recognition contrast would appear to represent a relatively clear instantiation of the RIC.

tency. That is, we examined the proportion of "high-frequency" judgments as a function of whether those judgments were made quickly or more slowly. The logic underlying this analysis is that the fastest responses should reflect more automatic, fluency-based processes, while slower responses should reflect more controlled processes (Hintzman & Curran, 1994; Toth, 1996; Yonelinas & Jacoby, 1994). This form of analysis allowed us to examine two separate, albeit related, issues with respect to normative frequency judgments. The first concerns the possibility of explicit contamination. If bias effects on frequency judgments do not reflect automatic influences of memory, but rather are based on more controlled (explicit) uses of memory, we should see a significant effect of prior presentation for a subject's slowest responses, but little or no effect on their fastest responses (cf. Ratcliff & McKoon, 1995).

The second issue addressed by the latency analyses concerns the possible operation of bias-correction processes in the making of normative judgments. That is, in contrast to the explicit contamination hypothesis, which proposes that explicit uses of memory would result in an artifactual increase in priming, evaluative judgment tasks may often reflect a fast, heuristic (fluency-based) process that is then *opposed* by a slower, more analytic (recollection-based) process. An example of such a pattern of fast bias followed by a slower correction has been shown by Hintzman and Curran (1994) in the context of recognition memory (see also McElree, Dolan, & Jacoby, 1999; Toth, 1996). Using a response-signal procedure to examine retrieval dynamics, they found that presentation of test lures which were highly similar to study items produced an initial, fast increase in false alarms, which were then suppressed by the slower accumulation of recalled information. If a similar process of bias followed by bias correction was to occur in the present judgment task, we should expect to see positive priming effects for subjects' fastest judgments, but little or no effect on their slower judgments (a pattern opposite to that predicted by the explicit contamination hypothesis). Indeed, assuming that subjects are highly sensitive to the biasing effects of prior

presentation, one might even predict a negative (i.e., below baseline or "contrast") effect on the slowest judgments, reflecting an overcorrection of the bias effect (see Herr, Sherman, & Fazio, 1983; Lombardi, Higgins, & Bargh, 1987; Martin & Achee, 1992; Thompson, Roman, Moskowitz, Chaiken, & Bargh, 1994).

EXPERIMENT 1a: ATTENTION AND STUDY REPETITION

The first experiment was designed simply to determine whether prior presentation of a word could increase the probability that the word would later be judged as having a high frequency of occurrence in the English language. At encoding, participants were shown a single list of words, some of which were presented once, others twice. Half of the subjects encoded the words under full attention and intentional memory instructions, whereas the other half encoded the words under incidental memory instructions while performing a demanding secondary task.² Immediately following encoding, participants were given the frequency judgment task which was followed by a structured interview that assessed the subjects' strategic approach to the task.

We manipulated item repetition at study because prior research has shown that while repeating items usually increases explicit test performance, repetition may have little or no effect on implicit memory performance (Challis & Sidhu, 1993; Roediger & Challis, 1992; Weldon, Roediger, Beitel, & Johnston, 1995). However, more recent work by Jacoby (1999) has suggested that study repetition can increase automatic influences of memory (i.e., familiarity) in recognition memory tasks. Similarly, research in the social literature has shown that the priming of trait judgments (see Higgins, Bargh,

²Following other researchers (e.g., Jacoby, 1991) we intentionally confounded attention condition (Full vs. Divided) with encoding instructions (intentional vs. incidental) in order to make the encoding tasks appear reasonable to the subjects. This seemed especially relevant for the Divided Attention subjects for whom we wanted minimal attention paid to the words. Intentional instructions would have undermined this goal. Attention condition and encoding instructions were unconfounded in Experiment 1b.

& Lombardi, 1985; Srull & Wyer, 1979) and ability-related performances (Dijksterhuis & Knippenberg, 1998) can be enhanced using intense or prolonged primes. Given the similarity of frequency judgments to recognition-memory judgments (i.e., both involve an intact word presented at test, and both may be influenced by processing fluency) and their similarity to trait judgments (both involve a subjective evaluation) we thought it important to assess the effects of study repetition.

More important was the question of whether priming would be affected by the degree of attention paid to the words at encoding. Prior research has shown that performance on *perceptual* implicit tests, such as stem completion and word identification, is insensitive to manipulations of attention (Schmitter-Edgecombe, 1996; Mulligan, 1998). However, research using *conceptual* implicit tests, such as exemplar generation or general-knowledge questions, has shown decreases in implicit memory under conditions of divided attention (Schmitter-Edgecombe, 1996; Mulligan, 1997, 1998). Manipulating attention may therefore provide insight into the informational basis of any priming effects on frequency judgments.

Method

Participants and design. Twenty-six undergraduate students enrolled in psychology courses at the Georgia Institute of Technology participated in return for course credit. The experiment was a 3×2 mixed design with study repetition [once presented ($1\times$), twice presented ($2\times$), and unstudied (new)] as the within-subjects factor and attention at encoding (Full vs. Divided) as the between-subjects factor. Based on responses in the strategy interview, described below, 2 subjects were replaced for judging frequency in the study list rather than frequency in the language, thus resulting in 24 participants (12 male), with a mean age of 19.8 ($SD = 1.2$).

Materials and counterbalancing. Ninety words, nouns and adjectives from 6 to 8 letters long, were used as critical items. Frequency of occurrence in the English language ranged from 18 to 35 per million, with an overall mean of 25.44 ($SD = 0.11$) as reported by Kučera and

Francis (1967).³ The 90 critical words were divided into three sets of 30 words each (A, B, and C). These three sets were equated for mean frequency (range, 25.27 to 25.53), word length (mean range from 6.73 to 7.13 letters), and, as well as possible, number of words having the same first letter. Ten additional words, similar to the critical items, were used as primacy and recency buffers at study, 5 words presented at the beginning and 5 at the end of each encoding list.

Participants were exposed to two of the three sets at study (e.g., A and B), with the third set (e.g., C) serving as new items on the frequency tests. Sets were rotated through conditions such that each set (and thus all words) served equally often in the $1\times$, $2\times$, and unstudied conditions. Three study lists were constructed for each combination of item sets (i.e., A&B, A&C, and B&C). For each list, once presented ($1\times$) items were randomly distributed throughout the list. Twice presented ($2\times$) items were also randomly distributed throughout the list, with the constraint that the second presentation of any item occurred in the second half of the list and had at least four other items presented between its first and second occurrence.

For the Divided Attention task, a tape-recorded series of digits from 0 to 9 was prepared, with the digits spoken at a 2-s rate. The subjects' task was to detect sequences of three odd digits. The digits were randomly generated with the constraint that no more than three odd digits could occur in sequence and that at least one, but no more than five numbers occurred between target sequences.

Equipment and procedure. For all experiments reported, stimuli were presented on IBM-compatible desktop computers using 14-in.

³The use of frequency norms from 1967 is, perhaps, less than optimal as it is likely that exposure frequencies change over time. It is worth noting, however, that more modern norms would not guarantee complete accuracy as it is also likely that exposure frequencies differ in different regions of the country (e.g., north vs. south) and for different participant populations (e.g., young vs. older adults). More importantly, all words were completely counterbalanced across conditions, thus eliminating any effects due to inaccuracies in the estimation of "true" frequency in the language.

color monitors. Words were presented in white against a black background using Turbo Pascal (v.5.0) sans serif font at default size 4; thus, each letter was at least 7×7 mm. Digits for the Divided Attention condition were presented via a portable tape recorder. All participants were run individually with an experimenter present.

The experiment began with participants filling out an IRB-approved consent form and a general demographic questionnaire. They were then told that they would be performing a series of tasks and that instructions would be given prior to each task. Instructions for all tasks (study and test) were provided on the computer screen. After reading the instructions for any task, participants were asked to describe the task they were about to perform; the experimenter then corrected any misunderstandings and reiterated the main parts of the task.

For the Full Attention condition, participants were told they would be seeing a list of words presented in the middle of the computer screen. Their task was to read each word aloud and to remember as many as possible for a later unspecified memory test. For the Divided Attention condition, participants were told they would be performing two tasks at once, a reading task and a digit monitoring task. They were told that they would be seeing a list of words presented on the computer screen which they were to read aloud. At the same time, they would be hearing a list of digits which they were to monitor for sequences of three odd digits. Subjects were told that, although it was important for them to read each word aloud, monitoring the digits was their primary ("more important") task. Subjects were instructed to press the space bar whenever they detected a target sequence. Pressing the space bar produced a short "beep" from the computer. Finally, they were told that should they miss a target sequence, the experimenter would say "miss" and that this was a signal to pay more attention to the digits task.

For both groups, the study list was preceded by a practice list of 10 words. This was done to give subjects in the Divided Attention condition practice in reading and monitoring. It was done in the Full Attention group to equate their expo-

sure to words with the Divided Attention group. The actual study list consisted of 100 trials (30 words presented once, 30 words presented twice, and 10 buffer words). Words were presented in the center of the computer screen, one at a time, for 1 s, followed by 1 s in which the screen was blank.

Immediately following the encoding task, subjects were given instructions for the frequency-judgment task. They were told that words would be presented on the computer screen and that their task was to judge whether each word was of high or low frequency in the English language. If a participant asked what was meant by "high" and "low" frequency, they were told that these terms referred to whether a word was "used often" or "used less often" in the English language. No examples were given. Participants made their responses by pressing the "/" key for a high-frequency response and the "z" key for a low-frequency response. Labels for this mapping, provided at the bottom of the screen, remained in view throughout the test.

The instructions also informed subjects that some of the words on the frequency test had also appeared on the prior (encoding) list. This overlap in stimuli was explained as simply due to the fact that all of our stimuli were drawn from the same pool and thus was irrelevant to their judgment task. In actuality, we informed subjects about the presence of studied words in order to avoid having them spontaneously "catch on" to the purpose of the experiment and thus intentionally rate studied words as either high or low frequency (see Bowers & Schacter, 1990).

The frequency-judgment test consisted of 90 words (30 $1\times$, 30 $2\times$, and 30 new) presented one at a time in a different randomized order for each participant. Words were presented in the center of the computer screen and remained in view until a response was detected. Detection of a response erased the word and then, following a 1-s delay, the next word was presented.

Immediately following the frequency test, subjects were given a structured interview concerning their approach to the task. In particular, the experimenter asked the following questions: (1) *What do you think was the purpose of the*

frequency-judgment task? (2) Did you use any particular strategy to make your frequency judgments? (3) During the frequency task, did you notice any consistent difference between the words you were asked to rate? (4) As mentioned, there were some words presented in the frequency-judgment task that were also presented in the initial study list. What percentage of trials in the frequency-judgment task do you think contained a word from the initial study list? (5) Did the presence of words from the initial study list interfere with your ability to make frequency decisions? (6) When making your frequency decisions, did you ever intentionally rate a word as high or low frequency because you remembered it as occurring on the initial study list? Following the interview, subjects were debriefed, given a course credit slip, and excused. The entire procedure took about 30 min.

Results and Discussion

The probability of accurately detecting a target sequence for the digit monitoring task was .84. This detection rate is comparable to those obtained in other studies using a similar listening task (e.g., Jacoby, Toth, & Yonelinas, 1993).

Performance on the frequency-judgment task is presented in Fig. 1 as a function of attention at encoding (Full vs. Divided) and study repetitions

(1×, 2×, and new). Note that for this and subsequent figures, error bars are 95% within-subjects confidence intervals (Loftus & Masson, 1994). As can be seen in the figure, prior presentation increased the proportion of high-frequency judgments but only in the Full Attention (FA) condition; subjects in the Divided Attention (DA) condition showed no effect of the prior study experience. These impressions were supported by an analysis of variance (ANOVA) which, although producing no main effects of attention, $F < 1$, demonstrated a significant effect of item repetition, $F(2,44) = 4.42$, $MSe = .005$, $p < .02$, as well as a significant interaction between the two factors, $F(2,44) = 3.81$, $MSe = .005$, $p < .03$. Planned comparisons showed that subjects in the FA group made significantly more “high-frequency” responses to the 1× and 2× items than to the unstudied items [t 's(22) = 3.85 and 4.11, respectively; p 's < .005]. The small difference between 1× and 2× was not significant ($t < 1$), nor were any of the differences between items in the DA group ($F < 1$).

The failure to find priming effects in the DA condition could be due to at least two factors. One possibility is that the priming of frequency judgments reflects implicit memory for prior conceptual (meaning-based) processing. As noted above, prior research has shown that

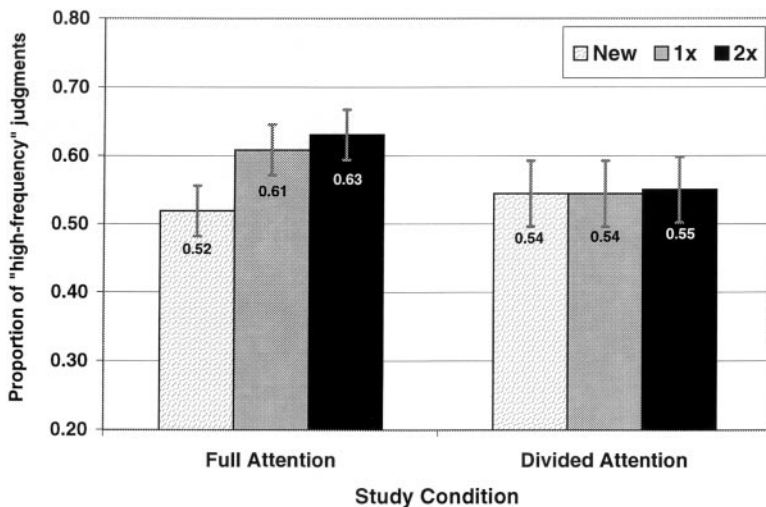


FIG. 1. “High-Frequency” Judgments as a function of Attention and Study Repetitions in Experiment 1a.

priming on conceptual implicit tests is reduced or eliminated by DA at encoding (e.g., Schmitter-Edgecombe, 1996; Mulligan, 1997, 1998). A second possibility is that, although ostensibly implicit (given our test instructions), the frequency test was contaminated by conscious/explicit uses of memory. As shown in prior research (Bowers & Schacter, 1990), if subjects suspect a relation between an implicit test and a prior encoding experience, they may use explicit memory to enhance their implicit test performance. In the present case, this would mean that subjects in the FA condition were judging words as “high frequency” based on explicit memory for these words. Divided attention, in contrast, would be expected to reduce explicit memory for the words’ prior presentation, thus eliminating any “priming” effects for this group.

Some evidence against the explicit-contamination hypothesis is provided by responses to our posttest interview. In particular, when asked to give the percentage of studied words on the frequency-judgment tests (see Question 4 under Methods), subjects significantly underestimated [mean estimate of 35.5% vs. the true 67%; $t(23) = 7.92, p < .001$]. Moreover, when subjects were asked if they had ever rated a word as high or low frequency because they remembered it being on the encoding list, only 2 of 26 subjects responded positively (both in the FA condition) and these subjects were eliminated from the data analysis. Indeed, most of the other subjects expressed genuine surprise when told during debriefing that the purpose of the study list was to influence their frequency judgments. These data, although not unassailable, are consistent with the claim that subjects were not using explicit memory to make their frequency judgments.

To investigate the issue further, we examined judgment performance as a function of response latency. For each subject, we first determined the median response time (RT) to unstudied words. This RT was then used to separate items in the two study conditions (1× and 2×) into “faster” and “slower” response subsets for which the proportion of high-frequency judgments was then recalculated. As described above, faster responses should better reflect

more automatic, fluency-based processes, while slower responses should reflect more consciously controlled recollection (Hintzman & Curran, 1994; Toth, 1996). Thus, if the priming effects observed in the FA condition were not due to automatic/implicit influences of memory, but rather reflected explicit uses of memory, we should expect to see a significant effect of prior presentation for the slower responses, but little or no effect for the faster responses.

The mean of the median RTs to unstudied words was 1228 ms for the FA group and 1079 ms for the DA group ($t < 1$). Performance on the frequency-judgment task as a function of study repetition and speed of response is presented in Fig. 2a for the FA group and Fig. 2b for the DA group. Looking first at FA, one can see that, contrary to the explicit contamination hypothesis, fast responses resulted in priming effects that were as large as those observed for the slower responses. This was confirmed by an ANOVA which showed a main effect of response speed, $F(1,11) = 10.60, MSe = .028, p < .01$, and study repetition, $F(2,22) = 7.59, MSe = .011, p < .005$, but no interaction, $F < 1$. These results thus suggest that the priming effect obtained in the overall data was not due to subjects intentionally responding “high frequency” to words they recognized from the prior study list. Indeed, the slower, presumably more controlled, responses for the 2× items showed a slight downward trend in high-frequency judgments.

In contrast to FA, subjects in the DA group (Fig. 2b), while showing a main effect of response speed [$F(1,11) = 34.81, MSe = .015, p < .001$], showed neither a main effect of repetition nor a repetition-by-speed interaction [both F 's < 1.42]. Posthoc analyses comparing only the 2× and unstudied conditions did show reliable priming for fast responses, $t(11) = 2.19, p < .03$, but based on the relatively small N , we hesitated to make much of this result, pending replication in Experiment 1b. The small “negative priming” or contrast effect hinted at in the slow responses failed to reach significance, $t(11) = 1.13, p = .282$.

In summary, judgments of normative frequency were positively biased when to-be-

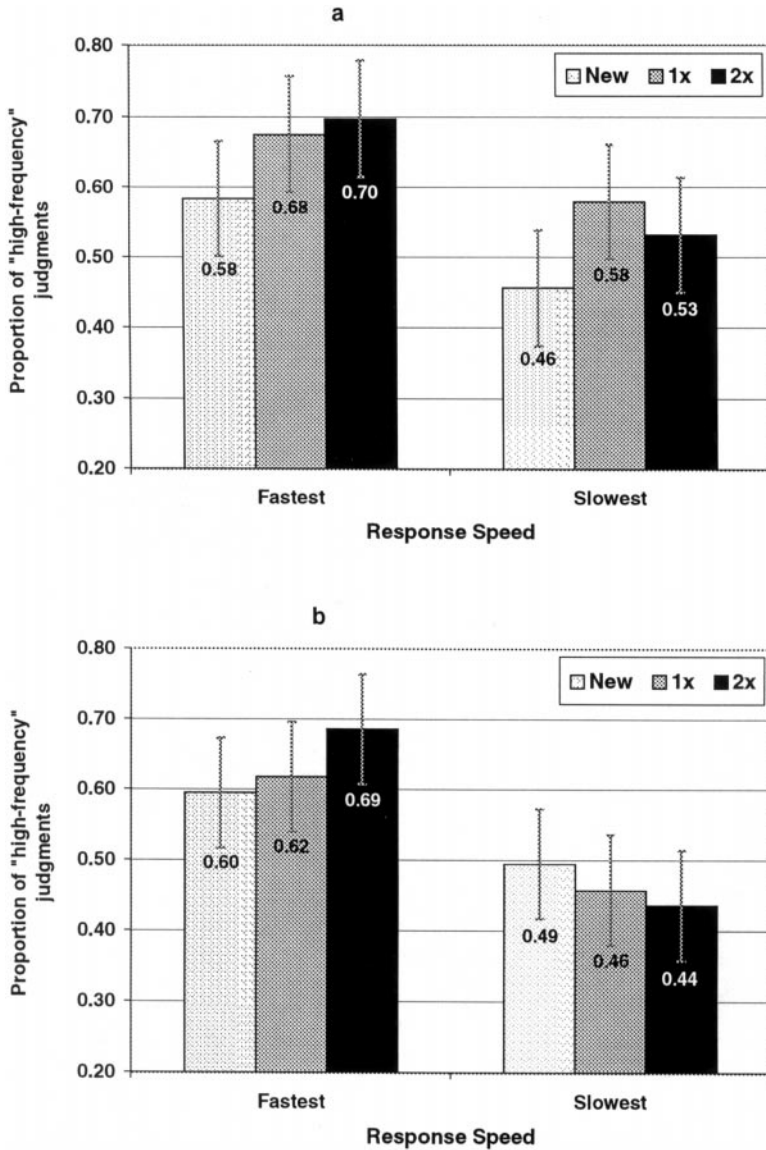


FIG. 2. “High-Frequency” Judgments as a function of Study Repetitions and Response Speed in Experiment 1a. (a) Full Attention at Study; (b) Divided Attention at Study.

judged words were previously encoded under conditions of full attention, but not when they were encoded under conditions of divided attention. Moreover, data from both posttest interviews and posthoc latency analyses suggested that the effects observed in the FA condition were not due to conscious, explicit uses of memory, but were indeed implicit (unconscious and unintentional). We comment on these find-

ings after first presenting a second experiment, designed to replicate the current results.

EXPERIMENT 1b: ATTENTION, STUDY REPETITION, AND COMPARISON TO RECOGNITION MEMORY

Experiment 1b was conducted as a conceptual replication of Experiment 1a. As in the first experiment, study words were presented either

once or twice, with half of the subjects encoding these words under conditions of FA and the other half encoding the words while performing a secondary task (DA). Unlike the first experiment, attention condition and study instructions were not confounded; that is, subjects in both attention conditions were told to intentionally study the words for a later memory test. The other major change to Experiment 1b was the inclusion of a recognition memory test which followed the frequency test.

Method

Participants and design. Thirty-six undergraduate students (24 male; mean age, 19.9; $SD = 2.0$), from the same source as in Experiment 1a, participated in return for course credit. The experiment was a $3 \times 2 \times 2$ mixed design with study repetition ($1\times$, $2\times$, and new) and test type (word frequency and recognition memory) as within-subjects factors and attention at encoding (Full vs. Divided) as the between-subjects factor.

Materials and counterbalancing. One-hundred and twenty words, nouns and adjectives from six to eight letters long, were used as critical items. Frequency of occurrence in the English language ranged from 18 to 35 per million, with an overall mean of 25.13 ($SD = 5.47$), as reported by Kučera and Francis (1967). The 120 critical words were divided into three sets (A, B, and C), each of which was further divided into two subsets (1 and 2). These six subsets were equated for mean frequency (range, 25.10 to 25.15), word length (mean range from 6.80 to 6.95), and, as well as possible, number of words having the same first letter. Ten additional words, similar to the critical items, were used as primacy and recency buffers at study, 5 words presented at the beginning and 5 at the end of each encoding list.

Participants were exposed to two of the three sets at study (e.g., A and B), with the third set (e.g., C) serving as new items on the frequency and recognition tests. The method for constructing the lists was identical to that used in Experiment 1a. Sets were rotated through conditions such that each set (and thus all words) served equally often in the $1\times$, $2\times$, and unstudied con-

ditions. For each set of words, half (e.g., subset A1) were presented on the frequency-judgment task, while the other half (e.g., A2) were presented on the recognition-memory test. Distribution of subsets to the two tests was also rotated such that each subset (and thus all words) served equally often in the frequency and recognition tests. The audiotape used for the DA task was the same as that used in Experiment 1a.

Procedure. The general methods for instructing subjects and presenting stimuli were identical to those used in Experiment 1a. The only changes were as follows. First, subjects in the DA condition were given intentional study instructions. That is, they were told the DA task required them to “do two things at once: Monitor a series of digits while simultaneously reading and trying to remember a list of words.” They were further told that, “although it is important for you to read the words aloud and to remember as many as possible . . . the number-monitoring task is more important.” The second change was that, for the frequency-judgment task, all subjects were instructed to respond accurately, but quickly, “not thinking too much about each word.” The final change was that, immediately following the frequency-judgment task, all subjects were given a recognition memory task for half of the study words (i.e., the half not presented on the frequency test). The recognition instructions told subjects that words would be presented on the computer screen and that their task was to judge whether these words had been presented on the earlier study list. Instructions as to response speed were not specified, but if participants asked, they were told that accuracy was now more important than speed. Participants made their responses by pressing the “/” key for a “yes” response and the “z” key for a “no” response. Labels for this mapping were provided at the bottom of the screen and remained in view throughout the test. The frequency and recognition-memory tests each consisted of 60 unique words (20 $1\times$, 20 $2\times$, and 20 new) presented one at a time in a different randomized order for each subject. Immediately following the recognition task, subjects were given the same strategy interview used in Experiment 1a.

Results and Discussion

The probability of accurately detecting a target sequence for the digit monitoring task was .83. This detection rate is comparable to that found in Experiment 1a and to that obtained in previous studies (e.g., Jacoby et al., 1993).

Performance on both the frequency-judgment and recognition-memory tests is presented in Fig. 3 as a function of attention at encoding (Full, Divided) and study repetition (1×, 2×, and new). Analysis of the frequency-judgment data revealed only a significant main effect of repetition, $F(2,68) = 3.71$, $MSe = .0111$, $p < .05$. The attention-by-repetition interaction was not reliable, $F < 1$, thus failing to replicate the reliable interaction found in Experiment 1a. We believe that the difference in statistical results mainly reflects the use of intentional study instructions for the DA group in Experiment 1b. These instructions likely resulted in subjects paying more attention to the study words (relative to Experiment 1a), thus producing a slight increase in priming as well as increased variability. Despite these differences, planned comparisons showed that the basic pattern of priming from Experiment 1a was replicated: For the FA group, the proportion of high-frequency judgments was reliably above baseline for both the 1× and 2× items [t 's (17) = 2.26 and 2.10, respectively, both p 's $< .05$]. In contrast, high-

frequency judgments did not differ among item types in the DA group (all t 's < 1.10).

Analysis of the recognition data revealed main effects of both attention, $F(1,34) = 14.34$, $MSe = .0330$, $p < .005$, and repetition, $F(2,68) = 206.41$, $MSe = .0142$, $p < .001$. These data replicate previous findings in the literature (for attention, see Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; Jacoby, 1991; for repetition, see Weldon et al., 1995). As well, attention and repetition interacted, $F(2,68) = 19.72$, $MSe = .0147$, $p < .001$. The interaction reflects the "mirror effect" (Glanzer & Adams, 1993) seen in Fig. 3. That is, relative to the FA condition, DA both decreased hit rates for studied words and increased false alarm rates for unstudied words. Most important for our purposes, however, is the main effect of study repetition. As can be seen, while repetition had a large effect on recognition memory, it had no observable effect on frequency judgments. This dissociation replicates previous findings in the implicit test literature (Challis & Sidhu, 1993; Roediger & Challis, 1992; Weldon et al., 1995) and thus provides support for the claim that priming on the frequency-judgment test reflected implicit uses of memory. Additional support was provided by subjects' estimation of the percentage of old words on the frequency test: As in Experiment 1a, although the true percentage of old words

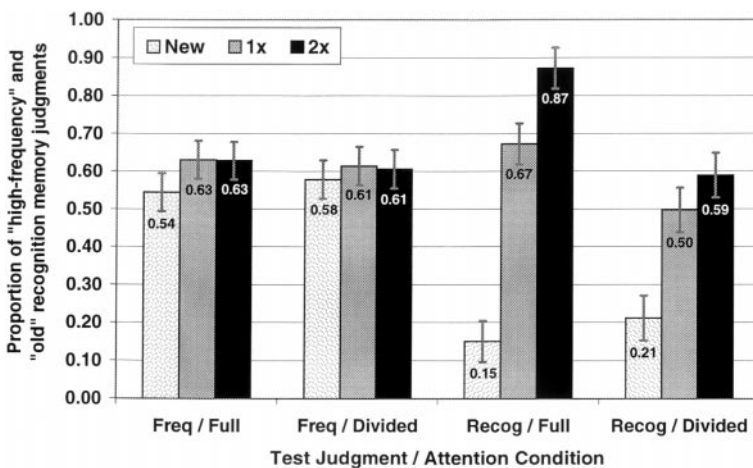


FIG. 3. Frequency and Recognition Judgments as a function of Attention and Study Repetitions in Experiment 1b.

was 67%, subjects significantly underestimated this value [mean estimate of 27.5%; $t(35) = 15.33, p < .001$]. Moreover, when subjects were asked if they had ever rated a word as high or low frequency because they remembered it being presented on the study list, none of the 36 subjects responded positively.

As in Experiment 1a, we also examined frequency judgments as a function of response latency. The mean of the median RTs to unstudied

words was 926 ms for the FA group and 918 ms for the DA group ($t < 1$). Performance on the frequency task as a function of study repetition and response speed is presented in Fig. 4a for the FA group and Fig. 4b for the DA group. As can be seen, these data also replicated the basic patterns observed in Experiment 1a. Analysis of the FA data revealed a main effect of response speed, $F(1,17) = 8.02, MSe = .0306, p < .05$, and a marginally reliable main effect of study

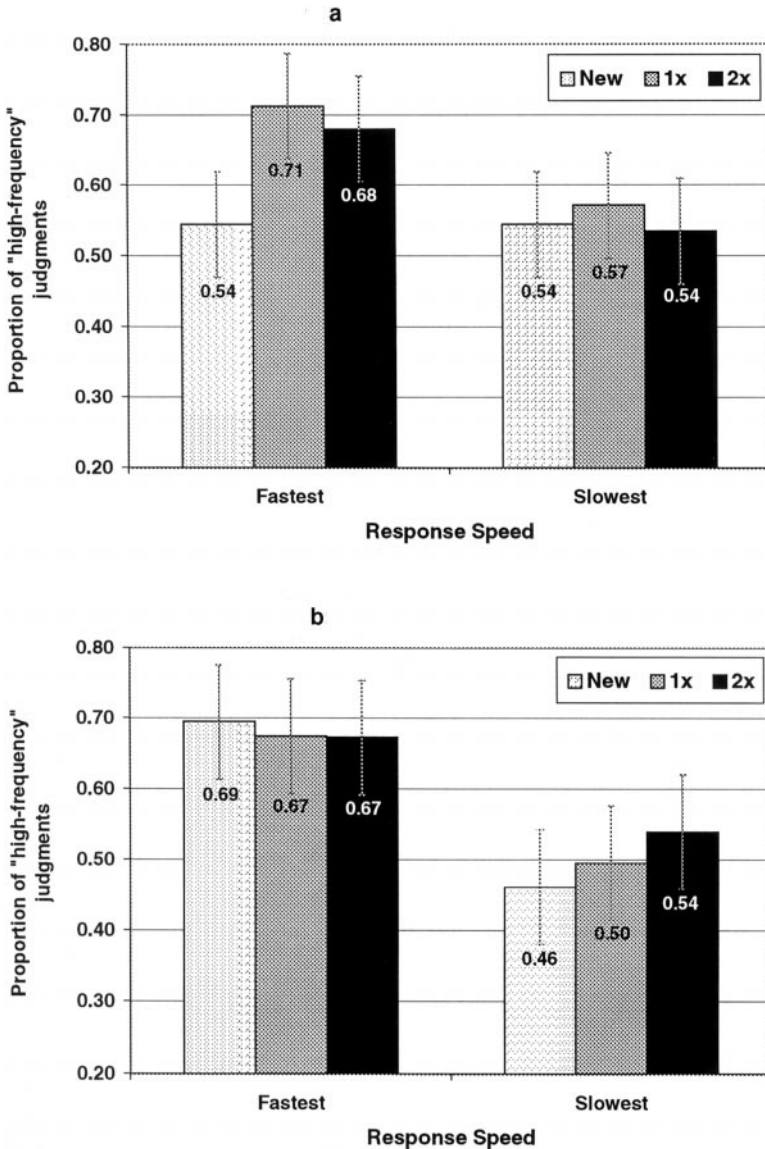


FIG. 4. "High-Frequency" Judgments as a function of Study Repetitions and Response Speed in Experiment 1b. (a) Full Attention at Study; (b) Divided Attention at Study.

repetition, $F(2,34) = 2.84$, $MSe = .0308$, $p = .072$. More importantly, the interaction between repetition and response speed was significant, $F(2,34) = 3.54$, $MSe = .0173$, $p < .05$. The interaction reflects the fact that while the fastest responses produced highly reliable, but not differential, priming for both $1 \times [t(17) = 3.92$, $p < .005]$ and $2 \times$ items [$t(17) = 2.99$, $p < .01$], the slowest responses showed no reliable priming (all t 's < 1). Note that these data are contrary to the explicit contamination hypothesis and instead suggest that, rather than elevating the observed proportion of high-frequency judgments, the slower responses operated to reduce, or correct for, the bias that occurred with the faster responses.

In contrast to the FA data, subjects in the DA group (Fig. 4b), while showing a main effect of response speed [$F(1,17) = 17.64$, $MSe = .0507$, $p < .001$], showed neither a main effect of repetition nor a repetition-by-speed interaction [F 's < 1]. Planned comparisons agreed with these analyses, showing no hint of priming for either the fastest or slowest responses [all t 's < 1.1].

In summary, Experiment 1b replicated the major findings of Experiment 1a in showing judgments of normative frequency to be positively biased when to-be-judged words were encoded under conditions of full attention, but not when they were encoded under conditions of divided attention. Moreover, data from posttest interviews and from posthoc latency analyses, as well as the dissociation between frequency and recognition judgments as a function of repetition, suggest that the effects observed in the FA conditions of both experiments were not due to conscious, explicit uses of memory, but were indeed implicit. Given prior research on implicit memory, this pattern of data suggests that judgments of normative word frequency are based on prior conceptual encoding of the priming words. Experiment 2 investigated this possibility further by manipulating level of processing at encoding.

EXPERIMENT 2: LEVELS OF PROCESSING

Experiment 2 had three major goals. The first was simply to replicate the priming effects ob-

served in the FA conditions of Experiments 1a and 1b. The second was to diagnose the failure to find bias effects in the DA conditions of those experiments. As discussed above, the two most obvious explanations for that failure are (a) contamination by explicit memory and (b) the possibility that, although not contaminated, the priming of frequency judgments reflects prior conceptual processing. Data from both the strategy interviews and latency analyses in Experiment 1 would appear to rule out any strong form of the explicit-contamination hypothesis. Experiment 2 therefore examined the conceptual processing notion by manipulating Level of Processing (LoP). A final goal of Experiment 2 was to rule out the possibility (suggested by a reviewer) that our results were an artifact of how responses (low vs. high frequency) were assigned to hands (left vs. right). To address this issue, the "high-frequency" response was assigned to the right hand for half of the subjects and to the left hand for the other half.

Method

Participants and design. Fifty-one undergraduates, from the same source as Experiment 1, participated in return for course credit. Three subjects were eliminated, 1 for intentionally rating words from the encoding phase as "high frequency," 1 for producing more false alarms than hits on the recognition test, and the third for a self-reported inability to consistently use the response keys on the recognition test. Thus, 48 subjects (24 male) participated in the experiment proper, with a mean age of 19.5 ($SD = 1.70$). The experiment was a $3 \times 2 \times 2$ mixed design with study processing (semantic, nonsemantic, and unstudied) and test type (frequency vs. recognition) as within-subjects factors and response hand assignment as the between-subjects factor. Twenty-four subjects made "low-frequency" judgments with their left hand and "high-frequency" ones with their right hand; the other 24 used the opposite assignment.

Materials and counterbalancing. The critical words were identical to those used in Experiment 1b, as was their separation into sets (A, B, and C) and subsets (A1, A2, B1, etc.). The counterbalance scheme was also identical such that

all words served equally often in the semantic, nonsemantic, and unstudied conditions and in the frequency and recognition tests. Sixteen additional words, similar to the critical items, were used as primacy and recency buffers at study, 4 words at the beginning and end of each of the two encoding lists. In addition to counterbalancing both the word sets and the response hand assignments, order of encoding tasks was also counterbalanced such that half of the participants performed the semantic task first, while the other half performed the nonsemantic task first. Overall then, 24 participants were required for one pass through the counterbalance scheme (3 study sets by 2 study task orders by 2 test item distributions by 2 response hand assignments) and we had 2 participants in each unique combination. Within these counterbalance constraints, words were presented in a different randomized order for each participant at both encoding and test.

Procedure. The general methods for presenting stimuli and instructing subjects were identical to those used in the previous experiments. For the semantic (pleasantness rating) encoding task, participants were told they would be seeing a list of words and that their task was to think about the meaning of each word in order to judge its pleasantness. Subjects were told that, although they should be accurate, we were most interested in their first impressions, so they should make their judgments quickly. This was done to reduce the possibility of ceiling effects on the recognition test. Pleasantness ratings were made on a scale of 1 to 5 and this scale ("1 (unpleasant) . . . (pleasant) 5") remained at the bottom of the screen throughout the task; subjects made vocal responses which were entered into the computer by the experimenter. For the nonsemantic (enclosed letter) task, participants were told that their task was to determine the number of enclosed letters in each word. Examples of enclosed ("p" and "e") and non-enclosed letters ("c" and "w") were provided in the instructions. The phrase "Number of enclosed letters?" remained at the bottom of the screen throughout the task. Participants made their responses vocally, which were entered into the computer by the experimenter. For both encoding tasks, 48 words (40 critical, 8

primacy/recency) were presented in the center of the screen, one at a time, and remained in view until a response was made. Detection of a response erased the word and then, following a 1-s delay, the next word was presented.

Immediately following the second encoding task, participants were given instructions for the frequency judgment task. These instructions were identical to those used in Experiment 1b. Participants made their responses with the computer keyboard, half pressing the "/" key for a high-frequency response and the "z" key for a low-frequency response and the other half using the opposite hand/key assignments. Appropriate labels for these mapping were provided at the bottom of the computer screen and remained in view throughout the test.

Immediately following the frequency task, subjects were given the recognition-memory test. Instructions for this test were identical to those used in Experiment 1b. Subjects previously using the "/" for a high-frequency response now used that key for a positive ("yes") recognition memory response and the "z" key for a "no" response; subjects previously using the "z" for a high-frequency response now used that key for a positive ("yes") recognition memory response and the "/" key for a "no" response. Labels for these mappings, provided at the bottom of the screen, remained in view throughout the test. For both the frequency and recognition tests, 60 words (a randomized presentation of 20 semantic, 20 nonsemantic, and 20 unstudied words) were presented in the center of the computer screen and remained in view until a response was detected. Detection of a response erased the word and then, following a 1-s delay, the next word was presented. Immediately following the recognition-memory test, subjects were given the strategy interview described in Experiment 1a. They were then debriefed, given credit, and excused.

Results and Discussion

Response hand assignment had little, if any, effect on the observed pattern of frequency responses; neither the main effect of hand assignment nor the hand-by-LoP interaction were reliable (F 's < 1). Performance on the recognition

test also failed to show a main effect of response hand assignment ($F < 1$). The hand-by-LoP interaction did approach significance, $F(2,92) = 2.67$, $MSe = .013$, $p = .075$, but the difference between the two hand assignments on the pattern of recognition data was subtle; that is, when “old” was assigned to the right hand, the means were .19, .46, and .87 for the unstudied, nonsemantic, and semantic conditions, respectively; the corresponding values for the left hand assignment were .16, .37, and .89. Given the nonsignificant effects, we collapsed over hand assignment in the following analyses.

The proportion of high-frequency judgments on the frequency task and the proportion of words accepted (“yes”) on the recognition memory test are presented in Fig. 5 as a function of study processing (semantic, nonsemantic, unstudied). As can be seen, prior presentation affected performance on both tests, but did so differentially, as a function of study processing. Performance on the recognition test showed a typical LoP effect [$F(2,94) = 459.62$, $MSe = .0132$, $p < .001$], with semantic study producing much higher recognition than nonsemantic study [$t(47) = 18.07$, $p < .001$] which, in turn, was greater than the false alarm rate [$t(47) = 28.87$, $p < .001$]. Analysis of the frequency task also showed an effect of study processing, $F(2,94) = 6.09$, $MSe = .0100$, $p < .005$. In contrast to recognition, however, planned compar-

isons showed that while high-frequency judgments for both semantic and nonsemantic items were reliably above the unstudied baseline [t 's(47) = 3.20 and 2.96, respectively, both p 's $< .01$], performance did not differ between the two encoding conditions, $t(47) = .293$.

Overall, then, LoP produced a dissociation between frequency and recognition judgments, thus providing further evidence that priming in the frequency task was implicit. This conclusion was also supported by responses in the strategy interview. Of the 51 subjects run, only 1 stated that they had made frequency judgments based on explicit memory for a word's prior presentation and this subject was eliminated from the final data set. Moreover, as in Experiments 1a and 1b, the remaining (nonintentional) subjects significantly underestimated the percentage of studied word on the frequency test [mean estimate of 39% vs. the true 67%, $t(47) = 8.98$, $p < .001$]. All of these data are consistent with the claim that subjects were not using explicit memory to make their frequency judgments.

Coupled with the attention results from Experiment 1, the finding that implicit frequency judgments were not affected by LoP presents something of a puzzle. Based on prior research with conceptual implicit tests, the elimination of priming effects following dividing attention (Experiment 1) would seem to imply that frequency judgments are conceptually based. If so, how-

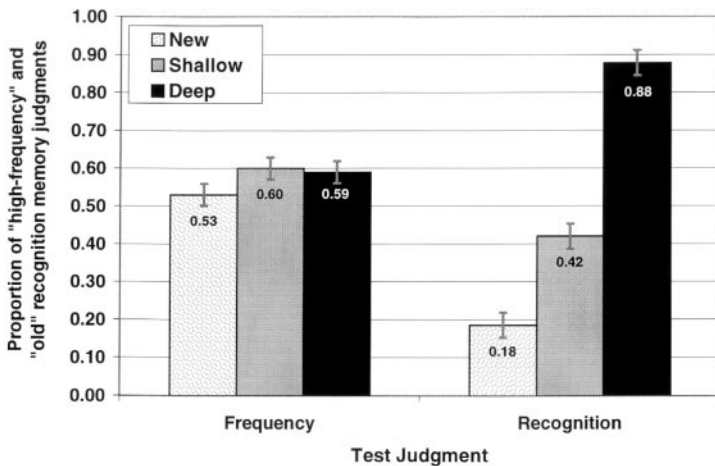


FIG. 5. Frequency and Recognition Judgments as a function of Level of Processing in Experiment 2.

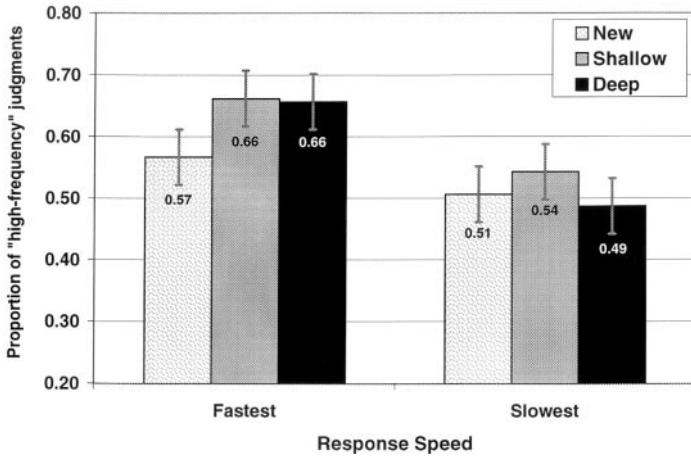


FIG. 6. "High-Frequency" Judgments as a function of Level of Processing and Response Speed in Experiment 2.

ever, one would expect an effect of LoP, but this did not occur. To investigate this issue further, we used the latency analysis employed in Experiments 1a and 1b. If, as suggested in those experiments, slower responses enabled subjects to counteract the high-frequency bias, perhaps there actually was an LoP effect on automatic retrieval, but this effect was hidden by the simultaneous presence of bias-correction processes.

Note that a latency analysis also has implications for the retrieval intentionality criterion (Schacter et al., 1989). If, as suggested by this criterion, our task was completely implicit, reflecting only unintentional influences of memory, then one might expect the pattern of performance to be roughly the same for the fastest and slowest responses as a function of study processing. In contrast, if the frequency-judgment task involved a mixture of controlled and automatic influences (consistent with the multiple-strategy perspective presented in the introduction), then one might expect to find an interaction between LoP and response speed.

Responses on the frequency task were separated into faster and slower subsets using each subject's median RT to unstudied items as the cutoff. The mean of these medians was 1022 ms ($SD = 223$). The proportion of high-frequency judgments as a function of study processing and speed of response is presented in Fig. 6. An initial analysis of these data showed that hand as-

signment had little, if any, confounding influence on the observed pattern of responses. That is, an ANOVA showed the main effect of hand, as well as the hand-by-LoP and hand-by-speed effects, to be unreliable, F 's < 1.65 , p 's $> .20$. Only the three-way interaction of hand-by-speed-by-LoP approached significance, $F(2,92) = 2.39$, $MSe = .0221$, $p = .098$, but the pattern of data underlying this interaction revealed no consistent confound as a function of response hand.⁴ Given the nonsignificant effects, we collapsed over this factor in the following analyses.

As can be seen in Fig. 6, speed of response did in fact change the pattern of performance. An ANOVA of these data showed main effects of response speed, $F(1,47) = 22.58$, $MSe = .0431$, $p < .001$, and study processing, $F(2,94) = 5.26$, $MSe = .0200$, $p = .01$. More importantly, the interaction between these two factors was also reliable, $F(2,94) = 3.13$, $MSe = .0227$, $p < .05$. Posthoc analyses showed that, for the fastest responses, both semantic and nonsemantic study resulted in a significant increase in high-frequency judgments compared to new

⁴For those subjects in which the "high-frequency" response was assigned to the right hand, the respective means for the new, $1\times$, and $2\times$ items were .57, .66, and .68 for the faster responses and .50, .54, and .43 for the slower responses. For the lefthand group, the respective means were .57, .66, and .63 for the faster responses and .52, .54, and .55 for the slower responses.

items [t 's(47) = 2.85 and 3.52, respectively, both p 's < .01]; the difference between the two study conditions was not reliable, $t(47) = .174$. For the slowest responses, neither the semantic [$t(47) = .655$] nor the nonsemantic items [$t(47) = .655$] were significantly different from the unstudied baseline; however, performance on the semantic items was marginally *below* that for the nonsemantic items, $t(47) = 1.96$, $p = .056$. Note that the latter finding (a *reverse* LoP effect) provides evidence against that claim that we had insufficient reliability, sensitivity, or power to detect an LoP effect in the present task (cf. Buchner & Wippich, 2000; Ostergaard & Jernigan, 1993).⁵

In summary, frequency judgments were dissociated from recognition memory judgments as a function of study processing (LoP). Based on the retrieval intentionality criterion, this result would appear to establish priming in the frequency-judgment task as both implicit and insensitive to prior conceptual processing. However, closer examination of the frequency judgments as a function of response speed revealed a more complex picture. Faster responses, which likely reflect more automatic influences of memory, showed no statistical difference between semantic and nonsemantic items. In contrast, slower responses, arguably based on more controlled forms of memory, showed a reversed LoP effect with performance on semantic items significantly below that for nonsemantic items. This reverse LoP effect is interesting and, similar to the data from Experiment 1, suggests the presence of a bias-correction process in subjects' frequency judgments. That is, given sufficient time, semantic processing—generally associated with increased conscious recollection—allowed subjects to correct for the word's automatic influence on their frequency judgments.

⁵ Given the reversed LoP effect, and the fact that no prior research has examined priming of normative word-frequency judgments, it is unclear what effect size should be used for formal power analyses. However, prior research using conceptual implicit tests (based on Brown & Mitchell's 1994 review) has found LoP effects to be around .13 (in proportional terms). The power to detect an effect of this size in present task was thus over .99.

The existence of a bias-correction process, and the fact that this correction was greater for semantic items, provides a possible explanation for the puzzle of why DA eliminated priming in the frequency task in Experiments 1a and 1b, thus suggesting a conceptual basis for frequency judgments, but LoP failed to reveal any evidence for conceptually driven priming. In particular, prior research on recognition memory has shown that, although speeding response times has a greater influence on conscious recollection than on more automatic forms of retrieval, recollection is not entirely eliminated even with relatively fast responses (e.g., Hintzman & Curran, 1994; Toth, 1995). As well, such "fast recollection" has been shown to differ as a function of prior study processing—being greater for semantic than for nonsemantic study items (see Toth, 1996, Experiment 3). Thus, if a fast recollection-like process was operating in the present experiment, was greater for semantic items, and was associated with correction of the implicit bias, this may explain why an LoP effect was not observed in the present study.

EXPERIMENT 3: REPETITION AND RESPONSE SPEED

Experiments 1 and 2 provide evidence that judgments of normative word frequency can be implicitly biased by a single prior exposure. This evidence was found both in the overall data of the two experiments (Full Attention conditions of Experiments 1a and 1b; Semantic and Nonsemantic conditions of Experiment 2) and when performance was examined as a function of response speed. In particular, the fastest responses in both experiments showed a significant high-frequency bias, while the slowest responses showed either no bias, or a slight low-frequency bias (Semantic/Slow-response condition of Experiment 2). However, given that the separation of responses into fast and slow was posthoc and may potentially reflect item selection artifacts, Experiment 3 was performed to experimentally manipulate response speed. Thus, subjects performed either the frequency or recognition test under two different response sets. In the fast-response condition, they were told to make their judgments "quickly and auto-

matically"; although told to be accurate, they were instructed that we were interested in "decision making under pressure" and thus should give us their "first impression." For the slow-response condition, in contrast, subjects were told that the accuracy of their judgments was our primary interest and thus that they should make "slow, considered" judgments.⁶

We expected results from this experiment to mimic those from the posthoc latency analyses done in Experiments 1 and 2. That is, performance in the fast-response condition was expected to show positive frequency effects (studied words > unstudied words), whereas the slow-response condition was expected to show no effect or possibly even an opposite effect.

Method

Participants and design. Forty-nine undergraduate students, from the same source as the previous experiments, participated in return for course credit. Based on the strategy interview, 1 subject was replaced for confusing high- and low-frequency responses keys, thus resulting in 48 subjects (16 male) with a mean age of 19.2 ($SD = 1.4$). The experiment was a $3 \times 2 \times 2$ mixed design with study repetition (1 \times , 2 \times , and new) and response speed (fast and slow) as the within-subjects factors and test type (frequency vs. recognition) as the between-subjects factor. Twenty-four subjects performed the frequency test and 24 performed the recognition test.

Materials. The 120 critical words plus 10 buffers were identical to those used in Experiment 1b. These words were divided into three

sets (A, B, and C), each of which was further divided into two subsets to accommodate the fast- and slow-response conditions. As in the previous experiments, the distribution of subsets was rotated such that each subset (and thus all words) served equally often in the each of the study repetition and response speed conditions. The repetition manipulation at study was instantiated using the same scheme employed in Experiment 1a. In addition to counterbalancing the word sets, the order of test conditions was also counterbalanced such that half of the subjects performed the fast-response condition first, while the other half performed the slow-response condition first. Overall, then, 24 subjects were required for one pass through the counterbalance scheme [3 study sets by 2 tests by 2 response speed orders by 2 test item distributions] and we had two subjects in each unique combination.

Procedure. The methods for presenting stimuli and instructing subjects were identical to those used in the previous experiments. At encoding, subjects were told they would see a list of words and that their task was simply to read each word aloud and to remember as many as possible for a later unspecified test.

Immediately following encoding, participants were given instructions for the first test (either the fast- or slow-response condition, depending on the counterbalance). The instructions for the frequency and recognition tests were similar to those used in the prior experiments, but modified to incorporate the response speed manipulation. Subjects in the fast-response conditions were instructed to make their judgments quickly. Although encouraged to be accurate, they were told that the goal of the test was to simulate decision-making under time pressure and thus that they should "not think too much" about each word but rather "go with their first impression." In contrast, subjects in the slow-response conditions were told to "think carefully" about each word and to respond only when they were confident about their decision. They were told to make slow, deliberate judgments and to emphasize accuracy rather than speed. For both the fast and slow conditions, 60 words (20 1 \times , 20 2 \times , and 20 new) were presented in the same

⁶As noted by a reviewer, a more formal response-signal paradigm (e.g., Hintzman & Curran, 1994) would provide more detailed information on the time course of normative frequency judgments. We agree, but believe that instructional manipulations of response speed can also be informative because they take into account individual differences in processing speed and may also be more ecologically valid with respect to how evaluative judgments are made in real-world contexts. Moreover, a major purpose of Experiment 3 was to provide supportive evidence for the posthoc latency analyses performed in Experiments 1 and 2. We believe an instructional manipulation of response speed more closely corresponds to the task demands of those earlier experiments.

way as in the previous experiments. Immediately following the frequency test, subjects in this condition were given the strategy interview.

Results and Discussion

We first examined RTs to test items to determine whether subjects were complying with the fast/slow response instructions. The RT data, as a function of test, repetition, and instructions are presented in Table 1. Note that data from two subjects, one in the frequency test condition and one in the recognition test condition, were not used in the RT analyses because they produced RTs greater than 2 *SDs* above the mean in all of the slow-response conditions (1×, 2×, and new); their data were maintained in the proportional analyses below because, although long, their RTs showed the appropriate pattern with respect to the speed manipulation (i.e., the fastest judgments were made in the fast-response condition).

As can be seen in Table 1, RTs in the fast-response conditions were considerably faster than those in the slow-response conditions, and this was true at all levels of repetition and for every subject in both the frequency and recognition tests. Thus, our subjects were clearly following instructions. An ANOVA of the recognition RTs (using Greenhouse–Geisser adjusted probabilities) showed significant main effects of response speed, $F(1,22) = 84.49$, $MSe = 212997.53$, $p < .001$, and repetition, $F(2,44) = 8.04$, $MSe = 83042.79$, $p < .005$, as well as a reliable interac-

tion between the two, $F(2,44) = 5.69$, $MSe = 69986.01$, $p < .05$. The interaction appears to reflect the relatively large decrease in RTs from 1× to 2× in the slow-response condition. An ANOVA of the frequency test RT data (Greenhouse–Geisser adjusted) showed only a main effect of response speed, $F(1,22) = 94.95$, $MSe = 627715.22$, $p < .001$. Neither the main effect of repetition nor the repetition-by-speed interaction reached significance (F 's < 1.93 , p 's $> .17$).

The proportion of high-frequency judgments on the frequency task and the proportion of words accepted (“yes”) on the recognition test are presented in Fig. 7 as a function of response speed (fast, slow) and study condition (1×, 2×, new). Analysis of the recognition data revealed a main effect of repetition, $F(2,46) = 405.06$, $MSe = .0159$, $p < .001$, as well as a reliable repetition-by-speed interaction, $F(2,46) = 6.41$, $MSe = .0077$, $p < .005$. The interaction reflects the mirror effect that can be seen in Fig. 7. That is, in comparison to the slow-response condition, fast responding resulted in lower hit rates for both 1× and 2× items, but a higher false alarm rate.

Analysis of the frequency data also revealed a main effect of repetition, $F(2,46) = 6.184$, $MSe = .0107$, $p < .005$. As well, the repetition-by-speed interaction was just at the .05 significance level, $F(2,46) = 3.06$, $MSe = .0102$, $p = .057$, suggesting a differential effect of study repetition for the fast- and slow-response conditions.

TABLE 1

Response Latencies and Standard Deviations (in Milliseconds) as a Function of Response Instructions (Fast, Slow) and Study Repetitions (new, 1×, 2×) for the Frequency-Judgment and Recognition-Memory Tests in Experiment 3

	Study repetitions					
	New		1×		2×	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Frequency-judgment latencies (ms)						
Fast-response condition	832	206	822	201	811	195
Slow-response condition	2199	761	2192	844	2017	622
Recognition-memory latencies (ms)						
Fast-response condition	704	115	697	160	653	119
Slow-response condition	1572	550	1482	557	1166	361

Note. $N = 23$ for both test conditions (see text for details).

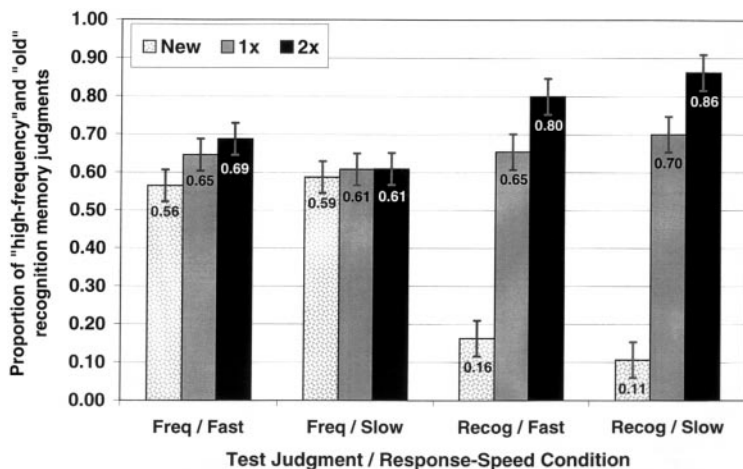


FIG. 7. Frequency and Recognition Judgments as a function of Response Instructions and Study Repetitions in Experiment 3.

This latter interpretation was supported by planned comparisons which showed that the proportion of high-frequency judgments for both the 1× and 2× items in the fast-response condition was significantly greater than baseline [$t(23) > 3.95$, p 's $< .001$]. The difference between 1× and 2× items in the fast-response condition was not reliable [$t(23) = 1.50$, $p = .148$], nor were any of the differences between item types in the slow response condition, all t 's ($23) < 1$.

In sum, results from the present response speed manipulation are consistent with those from our earlier posthoc analyses of response latencies: Fast responses—presumably based on more automatic, implicit processes—showed positive priming effects such that prior exposure to words increased the likelihood that those words were later judged as “high frequency.” In contrast, when slower, “more considered” judgments were rendered, prior exposure had no measurable influence on performance, suggesting the operation of a bias-correction process.

GENERAL DISCUSSION

The main finding of the present experiments is that a single prior exposure can increase the judged normative frequency of common English words. At the most general level, this finding could be taken as simply another demonstration

of implicit memory, of how performance can be influenced by prior experience in the absence of awareness or intent. However, we believe that three aspects of our experiments make their contribution more noteworthy. First, in comparison to studies examining repetition priming effects on identification or production tasks (see Gabrieli et al., 1999), many fewer studies have examined such effects on evaluative judgments, despite the fact that such judgments play a significant role in everyday life (see Tesser & Martin, 1996). Thus, similar to studies examining judgments of fame (Jacoby et al., 1989), truth (Begg et al., 1992), preference (Kunst-Wilson & Zajonc, 1980; Whittlesea, 1993), knowledge (Kelley & Lindsay, 1993), and personality (Srull & Wyer, 1979), the present studies help connect implicit memory research to the kind of evaluative judgments made in many real-world contexts (e.g., Menon, 1983).

A second contribution concerns the nature of the judgment itself. Prior studies examining repetition priming effects on evaluative judgments have typically employed novel stimuli that have no preexisting representations in memory. For example, both the false-fame and false-truth paradigms have employed novel or obscure stimuli which subjects have likely not previously encountered. Studies of trait priming in the social judgment literature have also tended to use novel

stimuli (e.g., the “Donald” paragraph) (e.g., Stapel, Koomen, & Zeelenberg, 1998; Thompson et al., 1994), as have studies investigating the priming of preference judgments (e.g., Korean melodies: Johnson, Kim, & Risse, 1985; abstract geometric shapes: Kunst-Wilson & Zajonc, 1980). In contrast, the current experiments examined judgments about stimuli (common English words) that have extensive preexisting representations in memory. As well, prior research has shown the judgment itself to be one at which subjects are quite good (Hasher & Zacks, 1984; Hintzman, 2000; Howell, 1973). This latter point seems especially important because, given extensive prior experience, subjects likely had available a number of different strategies for making their frequency judgments (Brown, 1995). Thus, demonstrating repetition priming effects on normative judgments is a theoretically important extension of prior research (see also Doshier & Rosedale, 1991; Menon, 1983; Srull, 1983; Wilson, Houston, Elting, & Brekke, 1996).

The third noteworthy contribution of the present experiments concerns the moderating effects of response latency. Experiment 3 found that priming in the frequency task occurred only for judgments rendered quickly; slower judgments failed to reveal any repetition effects. Importantly, this same speed-dependent pattern was also found in Experiments 1a, 1b, and 2, which did not directly manipulate response speed, but separated responses in a posthoc fashion. All of these findings point to a dynamic process of bias followed by bias correction that has been relatively unexplored in the unconscious influences and judgment literatures.

In the discussion that follows, we first describe theories that could potentially explain the positive priming effects observed in the present experiments, simultaneously asking how these theories would account for the elimination of priming for the slower responses (i.e., the issue of bias correction). We then relate our findings to recent theorizing on implicit memory.

Mechanisms Underlying Judgmental Bias and Bias Correction

The present studies make contact with a number of different research literatures, including

judgments of situational word frequency (e.g., Hintzman, 2000), judgments of behavioral frequency (Menon, 1983), and the nature of estimation strategies in general (Brown, 1995). Our results could also be related to formal models of memory, such as Hintzman’s (1988) MINERVA II which has been used to explicitly model frequency judgments. The present results also bear a striking similarity to assimilation and contrast phenomena as explored in the social judgment literature (e.g., Herr et al., 1983; Martin & Achee, 1992). Here, however, we restrict our comments to the fluency hypothesis and the source-monitoring framework, as these approaches have been most widely applied to judgment tasks similar to that used in the present experiments.

The fluency hypothesis states that experience with a stimulus, such as a word, produces changes in the component perceptual and conceptual processes such that, on subsequent encounters with that stimulus, processing is faster and more efficient (Jacoby & Dallas, 1981; Masson, 1989; Whittlesea, 1997). Such fluent reprocessing then acts as a source of information which—in combination with other sources, such as the goal of the judgment being performed—is used to make attributions about the target stimulus. Although the fluency hypothesis has largely been used to explain the subjective experience of familiarity for episodic events (Jacoby et al., 1989b; Whittlesea, 1993), it also provides a potential explanation for the priming obtained in the present experiments. Specifically, encountering words on the study list increased the fluency with which those words were later processed on the frequency-judgment test. Given the goal of assessing normative frequency, such fluent reprocessing resulted in heightened feelings of familiarity, thus increasing the probability that studied words were rated as high frequency.

Note that a critical component of the standard fluency-based explanation of judgmental bias (as applied to the false-fame task, for example) is the idea that familiarity stemming from a prior encounter is an undifferentiated or “sourceless” subjective experience; that is, the prior encounter enhances the item’s felt familiarity, but

does not specify the episodic details surrounding that encounter. An important question for this approach, then, is how subjects are able to avoid fluency-based biases. The most common mechanism appealed to is conscious recollection of episodic details. Thus, in the fame-judgment task, for example, recollection that a name was presented on a list of nonfamous names can allow subjects to avoid the false-fame effect (Jacoby et al., 1989; see also Ratcliff & McKoon, 1995). Similarly, recollection of episodic details has been shown to allow rejection of recognition test lures that are highly similar to studied items (Hintzman & Curran, 1994).

Importantly, however, such a *recollect-to-reject* strategy cannot provide a complete explanation of bias correction in the present task because recollection would not have specified the correct response. That is, unlike the fame paradigm (where recollection can unambiguously identify a name as nonfamous), and unlike recognition (where recollection can unambiguously identify a test word as a lure), recollection of a studied word in the present task would not have unambiguously identified that item as being low frequency. Even with perfect recollection of prior presentation (“*oh yes, I definitely remember 'jacket' being presented earlier*”), subjects would still be left with the judgment of the word’s preexperimental frequency (“*but how often is 'jacket' used in English?*”).

Two other ways in which recollection could act to reduce biases from a prior encounter have also been considered (see Kelley & Jacoby, 1996). First, recollection could trigger an adjustment process such that, after first making a heuristic assessment of an item’s familiarity, subjects then systematically adjust this familiarity downward for recollected items. Second, recollection could prompt subjects to forego familiarity altogether as a basis for judgment and instead employ more analytic judgment strategies (i.e., using their own “theory” of normative frequency). A limitation in applying these ideas to the current data, however, is that, in addition to consciously recollecting study items, both strategies require that subjects be aware of the potential biasing influence of prior presentations and be motivated to avoid such bias. As well, the

first *recollect-and-adjust* strategy would additionally require subjects to have some idea about the direction and magnitude of the bias (see Wilson & Brekke, 1994).

Against these possibilities, we found little evidence in our posttest interviews that subjects were attempting to recollect prior presentation on the study list or to counteract their influence. Thus, in response to Question 2 (“what strategy did you use to make your frequency judgments?”), no subject in any of the four experiments reported intentionally remembering prior presentations as a basis for judgment. And, as reported, subjects consistently underestimated the percentage of studied words on the frequency tests, suggesting that their focus was not on the relation between test items and the prior study list. We did find some evidence that subjects were aware that prior presentations could bias their judgments; thus, in response to Question 5 (“Did the presence of studied words on the test interfere with your ability to make frequency judgments?”), 31 of the 132 subjects (23.5%) noted that possibility. However, we believe that this value is likely inflated by the prior question (No. 4) which asked them to estimate the percentage of old words on the test (cf. Schwarz, 1999), thus making it unclear whether their answer to Question 5 reflected an “on-line” awareness of bias at the time of test or whether it reflected a retrospective assessment of the *possibility* of such bias. More importantly, only 6 of the 132 (4.5%) reported attempting to counteract a perceived bias. This low value is interesting, given the clear evidence for bias correction in all four experiments, but is not unprecedented; studies in the social judgment literature have also reported data consistent with judgmental correction in the absence of subjective reports of such correction (e.g., Martin, Seta, & Crelia, 1990; Thompson et al., 1994).

The present results could also be interpreted within Johnson’s source-monitoring (SM) framework (Johnson et al., 1993). Briefly, this framework is concerned with how people determine the origin of their memories, for example, whether a memory was based on internal/cognitive events or external/perceived events; whether external memories came from one or

another source (e.g., person A vs. B, television vs. radio, etc.); and on how such identifications influence, or are influenced by, judgment-related processes. Although much of the early research within the SM framework was directed at explicit source judgments, it has more recently been used to explain a number of implicit memory phenomena, including false memory (Mather, Henkel, & Johnson, 1997), eyewitness suggestibility (Lindsay, 1994), false fame (Dywan & Jacoby, 1990), and unconscious plagiarism (Marsh, Landau, & Hicks, 1997).

From an SM perspective, the subject's task in the present experiments was to judge the preexperimental familiarity of test items, while "gating-out" the familiarity associated with experimental presentations (cf. Raye, Johnson, & Taylor, 1980). Thus, when subjects could not recollect the experimental source of a test word, familiarity from that item's prior presentation contributed to their judgment and priming was observed. When subjects could remember a test item's prior presentation, they were able to successfully filter out familiarity from that presentation, base their judgment only on preexperimental familiarity, and no priming effects were observed. Note that, to the extent that such filtering depends on conscious recollection, this explanation is effectively identical to the fluency-based explanation described above and thus would have the same drawbacks. Thus, there was little evidence that subjects were engaged in a process of explicit source monitoring while performing the current task and little evidence that subjects believed their judgments to be biased, both of which would seem to be prerequisites for the filtering of inappropriate sources familiarity. However, if the reliance on conscious recollection is relaxed, then the SM characterization may provide more flexibility in explaining how subjects are able to avoid the effects of prior presentation on judgment. That is, similar to the criterion-shift argument that has recently been applied to the false-memory effect (Miller & Wolford, 1999), it seems possible that implicit retrieval of fragmentary details could affect the judgment criteria that subjects used to assess word frequency (see Raye et al., 1980).

Summarizing, both the fluency and SM approaches suggest that bias is a likely outcome of prior exposure and that correction of such bias will depend upon conscious recollection, or source identification, of previously presented items. What is not specified in these approaches is (a) how subjects are able to screen out familiarity from specific prior encounters; (b) how they adjust their responses (discounting familiarity) when filtering is not possible; and (c) the conditions under which they switch to more analytic judgment strategies, thus obviating the need for filtering or adjusting levels of familiarity. All of these seem to be critical issues for future research on judgmental biasing.

Autocorrection and the Effects of Preexperimental Experience on Implicit Memory

We have been discussing the elimination of bias in a subjects' slowest responses as due to an active, controlled process based on conscious recollection or source monitoring. Note that if that explanation is correct, then it would mean that our task was severely contaminated by explicit uses of memory. Moreover, it would suggest that such contamination can occur, even when standard criteria used to diagnose such contamination, such as self-reports, implicit/explicit dissociations, and the RIC, indicate the implicit test to be process pure. Although this is a serious charge, we want to quickly add that the processes operating in our *judgment* task appear importantly different from those operating on more standard implicit tasks measuring identification or production (such as naming, perceptual identification, or fragment completion). Specifically, explicit contamination of standard implicit tests usually results in an artifactual increase in measured priming. In contrast, explicit contamination of the present test (if indeed such contamination occurred) acted to decrease priming. In this sense then, our task appears to have functioned like a "natural" opposition or exclusion condition similar to that used in the process-dissociation procedure (Jacoby, 1991).

In contrast to the claim that our task was explicitly contaminated, however, another possibility is that subjects performed the task as in-

structed (i.e., implicitly, with respect to the study list), and thus that the elimination of the bias for the slower responses occurred automatically, simply as a function of the retrieval of preexperimental exposures. That is, one might speculate that a fundamental basis for judgments of normative frequency is the (likely unconscious and in-parallel) retrieval of all prior experiences with the to-be-judged stimulus, followed by a heuristic assessment of the number, strength, or global familiarity of these experiences (see Brown, 1995; Hintzman, 1988; Whittlesea, 1997). If one makes the further assumption that retrieval of multiple instances takes time, and that more recent (or more contextually similar) experiences are retrieved faster than more temporally dated (or more contextually dissimilar) experiences, then the present results could be explained without recourse to explicit uses of memory (i.e., recollection or source-monitoring mechanisms). On this *autocorrection* account, recent (or contextually similar) experiences with a to-be-judged event would tend to be accessed early during the judgment process, resulting in an overrepresentation of such experiences and thus judgmental bias. In contrast, delaying the judgment, either in response to task demands or the importance of the judgment, would allow more remote (or contextually dissimilar) experiences to be accessed, reducing the weight of the study experience and thereby producing an apparent correction of the fast bias.

Note the implications of this account. First, it would make the term “bias correction” something of a misnomer, at least in the sense that it has been treated as an intentionally controlled process (e.g., Wilson & Brekke, 1994). Second, it would account for the fact, noted above, that subjects expressed little if any explicit knowledge of the existence of a bias or of their attempts to correct for it. Finally, given that such an autocorrection process would be predicated on the existence of an established knowledge based about the to-be-judged stimuli, it would imply that judgments about well-known stimuli will be qualitatively different from judgments about novel stimuli—as, for example, has been studied with the false-fame, false-truth, and person/trait-priming paradigms. This point seems

especially relevant if one is interested in how unconscious influence manipulations would operate in real-world judgment contexts (cf. Herr et al., 1983; Srull, 1983; Wilson et al., 1996).

Note that our autocorrection account is also compatible with a recent view of priming put forth by Ostergaard (1998). Ostergaard argued that priming in any task should be interpreted in the context of the more general information processing demands of that task. By definition, priming reflects the influence of memory for a specific study experience. However, a subject's overall ability to perform an “implicit” task will likely draw not only on experimentally provided (“primed”) information, but also on information gained from preexperimental exposures to relevant stimuli. An interesting implication of this perspective is that “in some sense, priming effects may be regarded as ‘contamination’ of task performance by episodic information from the study episode” (Ostergaard, 1998, p. 56). This statement nicely captures the positive priming effects observed in the present experiments: Subjects were given a task that should have been (and likely was) predominantly performed via access to thousands of preexperimental exposures to the to-be-judged words. What is noteworthy is that a single, recent exposure was able to implicitly “contaminate” those judgments.

Ostergaard's (1998) perspective may also have implications for interpreting the effects of our experimental manipulations (i.e., study repetition and LoP). In particular, his Information Availability (IA) model proposes that the magnitude of priming effects on any particular task, as well as the effects of particular experimental manipulations, will be moderated by the amount of information available from other (e.g., preexperimental) sources contributing to performance. To demonstrate this point, he used a word naming task to show that manipulations of study repetition, word frequency, and study–test delay may have no observable influence on priming when overall naming performance is relatively high (obtained by instant-onset presentation of the to-be-named words, a method that presumably allowed preexperimental exposures to the words to play a large role in performance). In contrast, these same manipulations produced

highly significant effects when overall performance was made more difficult (obtained by fading-in of the to-be-named words, which presumably allowed the more recent study experience to exert a larger influence on performance).

These data, then, call into question the dissociations we obtained between frequency and recognition judgments as a function of study repetition and LoP. According to the IA model, repetition and LoP may have been having an influence on priming (i.e., the contribution of the study experience to performance), but those influences were effectively washed out by the larger contribution of preexperimental exposures to the judged words. We have no direct evidence against this interpretation and thus view the current frequency/recognition dissociations as only provisional. We do note, however, that should LoP be found to moderate the automatic influence of prior experience on frequency judgments, it would help resolve the anomalous findings with regard to divided attention (Experiment 1) and LoP (Experiment 2). And although this would mean that LoP was an associating, rather than dissociating, variable with respect to (implicit) frequency judgments and (explicit) recognition memory judgments, we believe there are other grounds for concluding that our subjects were not performing the frequency-judgment task as a test of explicit memory.

In summary, the present experiments examined the possibility that a single prior experience can produce automatic/implicit influences on judgments about a class of stimuli for which subjects have had extensive prior experience. Moreover, the judgment was one for which prior research has shown subjects to be quite accurate and to have available multiple strategies for judgment. The results showed that such priming effects can occur, but that they display complex temporal properties, occurring only for a subject's fastest responses. Perhaps most interestingly, the elimination of priming for the slower responses occurred in the absence of self-reports of awareness or intent to correct for bias. The studies thus suggest that implicit priming effects may be more complex than previously assumed when evaluative judgments are directed at a well-known class of stimuli.

REFERENCES

- Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 10, pp. 1–61). Mahwah, NJ: Erlbaum.
- Begg, I., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, **121**, 446–458.
- Begg, I., Maxwell, D., Mitterer, J. O., & Harris, G. (1986). Estimates of frequency: Attribute or attribution? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **12**, 496–508.
- Bowers, K. S. (1975). The psychology of subtle control: An attributional analysis of behavioural persistence. *Canadian Journal of Behavioural Sciences*, **7**, 78–95.
- Bowers, J. S., & Schacter, D. L. (1990). Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 404–416.
- Brown, A. S., & Mitchell, D. B. (1994). A reevaluation of semantic versus nonsemantic processing in implicit memory. *Memory & Cognition*, **22**, 533–541.
- Brown, N. R. (1995). Estimation strategies and the judgment of event frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 1539–1553.
- Brown, N. R., Rips, L. J., & Shevell, S. K. (1985). The subjective dates of natural events in very-long-term memory. *Cognitive Psychology*, **17**, 139–177.
- Brown, N. R., & Siegler, R. S. (1992). The role of availability in the estimation of national populations. *Memory & Cognition*, **20**, 406–412.
- Buchner, A., & Wippich, W. (2000). On the reliability of implicit and explicit memory measures. *Cognitive Psychology*, **40**, 227–259.
- Burton, S., & Blair, E. (1991). Task conditions, response formulation processes, and response accuracy for behavioral frequency questions in surveys. *Public Opinion Quarterly*, **55**, 50–79.
- Challis, B. H., & Sidhu, R. (1993). Dissociative effect of massed repetition on implicit and explicit measures of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **19**, 115–127.
- Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, **125**, 159–180.
- Dijksterhuis, A., & Knippenberg, A. v. (1998). The relation between perception and behavior, or how to win a game of Trivial Pursuit. *Journal of Personality and Social Psychology*, **74**, 865–877.
- Dywan, J., & Jacoby, L. L. (1990). Effects of aging on source monitoring: Differences in susceptibility to false fame. *Psychology and Aging*, **5**, 379–387.
- Gabrieli, J. D. E., Vaidya, C. J., Stone, M., Francis, W. S., Thompson-Schill, S. L., Fleishman, D. A., Tinklenberg, J. R., Yesavage, J. A., & Wilson, R. S. (1999). Convergent behavioral and neuropsychological evi-

- dence for a distinction between identification and production forms of repetition priming. *Journal of Experimental Psychology: General*, **128**, 479–498.
- Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 5–16.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, **108**, 356–388.
- Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist*, **39**, 1372–1388.
- Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, **19**, 323–340.
- Higgins, E. T., Bargh, J. A., & Lombardi, W. (1985). Nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **11**, 59–69.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, **95**, 528–551.
- Hintzman, D. L. (2000). Memory judgments. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 165–177). New York, NY: Oxford Univ. Press.
- Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, **88**, 297–306.
- Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language*, **33**, 1–18.
- Hock, H. S., & Hasher, L. (1990). Judgments of frequency: A tool for the analysis of memory. In J. W. Rohrbaugh, R. Parasuraman, & R. Johnson, Jr. (Eds.), *Event related brain potentials: Basic issues and applications* (pp. 158–164). New York: Oxford Univ. Press.
- Howell, W. C. (1973). Representation of frequency in memory. *Psychological Bulletin*, **80**, 44–53.
- Huppert, F. A., & Percy, M. (1978). The role of trace strength in recency and frequency judgements by amnesic and control subjects. *Quarterly Journal of Experimental Psychology*, **30**, 347–354.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, **30**, 513–541.
- Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25**, 3–22.
- Jacoby, L. L., Allan, L. G., Collins, J. C., & Larwill, L. K. (1988). Memory influences subjective experience: Noise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 240–247.
- Jacoby, L. L., & Brooks, L. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 18, pp. 1–47). New York: Academic Press.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306–340.
- Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989a). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, **56**, 326–338.
- Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989b). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Erlbaum.
- Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, **122**, 139–154.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3–28.
- Johnson, M. K., Kim, J. K., & Risse, G. (1985). Do alcoholic Korsakoff's patients syndrome patients acquire affective reactions? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **11**, 22–36.
- Jonides, J., & Jones, C. M. (1992). Direct coding for frequency of occurrence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **18**, 368–378.
- Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language, Special Issue: Illusions of Memory*, **35**, 157–175.
- Kelley, C. M., & Lindsay, D. S. (1996). Conscious and unconscious forms of memory. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 31–63). San Diego, CA: Academic Press.
- Kolers, P. A., & Roediger, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, **23**, 425–449.
- Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown Univ. Press.
- Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, **19**, 557–558.
- Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, **4**, 551–578.
- Lindsay, D. S. (1994). Memory source monitoring and eyewitness suggestibility. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 27–55). New York, NY: Cambridge Univ. Press.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review*, **1**, 476–490.

- Lombardi, W. J., Higgins, E. T., & Bargh, J. A. (1987). The role of consciousness in priming effects on categorization: Assimilation versus contrast as a function of awareness of the priming task. *Personality and Social Psychology Bulletin*, **13**, 411–429.
- Marsh, R. L., Landau, J. D., & Hicks, J. L. (1997). Contributions of inadequate source monitoring to unconscious plagiarism during idea generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23**, 886–897.
- Mather, M., Henkel, L. A., & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory & Cognition*, **25**, 826–837.
- McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25**, 563–582.
- Malthaup, K. S. (1995). Aging, source, and decision criteria: When false fame errors do and do not occur. *Psychology and Aging*, **10**, 492–497.
- Martin, L. L., & Achee, J. W. (1992). Beyond accessibility: The role of processing objectives in judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 195–216). Hillsdale, NJ: Erlbaum.
- Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology*, **59**, 27–37.
- Masson, M. E. J. (1989). Fluent reprocessing as an implicit expression of memory for experience. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit memory: Theoretical issues*. Hillsdale, NJ: Erlbaum.
- Means, B., & Loftus, E. F. (1991). When personal history repeats itself: Decomposing memories for recurring events. *Applied Cognitive Psychology*, **5**, 297–318.
- Menon, G. (1993). The effects of accessibility of information in memory on judgments of behavioral frequencies. *Journal of Consumer Research*, **20**, 431–440.
- Miller, M. B., & Woford, G. L. (1999). Theoretical commentary: The role of criterion shift in false memory. *Psychological Review*, **106**, 398–405.
- Mulligan, N. W. (1997). Attention and implicit memory tests: The effects of varying attentional load on conceptual priming. *Memory & Cognition*, **25**, 11–17.
- Mulligan, N. W. (1998). The role of attention during encoding in implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24**, 27–47.
- Ostergaard, A. L. (1998). The effects on priming of word frequency, number of repetitions, and delay depend on the magnitude of priming. *Memory & Cognition*, **26**, 40–60.
- Ostergaard, A. L., & Jernigan, T. L. (1993). Are word priming and explicit memory mediated by difference brain structures. In P. Graf & M. E. J. Masson (Eds.), *Implicit memory: New directions in cognition, development, and neuropsychology* (pp. 327–349). Hillsdale, NJ: Erlbaum.
- Ratcliff, R., & McKoon, G. (1995). Bias in the priming of object decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 754–767.
- Ratcliff, R., & McKoon, G. (1996). Bias effects in implicit memory tasks. *Journal of Experimental Psychology: General*, **125**, 403–421.
- Raye, C. L., Johnson, M. K., & Taylor, T. H. (1980). Is there something special about memory for internally generated information? *Memory & Cognition*, **8**, 141–148.
- Reingold, E. M., & Merikle, P. M. (1990). On the inter-relatedness of theory and measurement in the study on unconscious processes. *Mind and Language*, **5**, 9–28.
- Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology*, **39**, 475–543.
- Richardson-Klavehn, A., Lee, M. G., Joubbran, R., & Bjork, R. A. (1994). Intention and awareness in perceptual identification priming. *Memory & Cognition*, **22**, 293–312.
- Roediger, H. L., & Challis, B. H. (1992). Effects of exact repetition and conceptual repetition on free recall and primed word-fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **18**, 3–14.
- Roediger, H. L., & McDermott, K. B. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology* (Vol. 8, pp. 63–98). Amsterdam: Elsevier Science.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 501–518.
- Schacter, D. L., Bowers, J., & Booker, J. (1989). Intention, awareness, and implicit memory: The retrieval intentionality criteria. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit memory* (pp. 47–65). Hillsdale, NJ: Erlbaum.
- Schmitter-Edgecombe, M. (1996). The effects of divided attention on implicit and explicit memory performance. *Journal of the International Neuropsychological Society*, **2**, 111–125.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, **54**, 93–105.
- Seamon, J. G., McKenna, P. A., & Binder, N. (1998). The mere exposure effect is differentially sensitive to different judgment tasks. *Consciousness and Cognition*, **7**, 85–102.
- Srull, T. K. (1983). The role of prior knowledge in the acquisition, retention, and use of new information. In R. P. Bagozzi & A. M. Tybout (Eds.) *Advances in consumer research*, (Vol. 10, pp. 572–576). Ann Arbor, MI: Assoc. for Consumer Res.
- Srull, T. K., & Wyer, R. S., Jr. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology*, **37**, 1660–1672.

- Stapel, D. A., Koomen, W., & Zeelenberg, M. (1998). The impact of accuracy motivation on interpretation, comparison, and correction processes: Accuracy X knowledge accessibility effects. *Journal of Personality and Social Psychology*, **74**, 878–893.
- Tesser, A., & Martin, L. L. (1996). The psychology of evaluation. In E. T. Higgins & A. W. Kruglansky (Eds.) *Social psychology: Handbook of basic principles* (pp. 400–432). New York, NY: Guilford.
- Thompson, E. P., Roman, R. J., Moskowitz, G. B., Chaiken, S., & Bargh, J. A. (1994). Accuracy motivation attenuates covert priming: The systematic reprocessing of social information. *Journal of Personality and Social Psychology*, **66**, 474–489.
- Toth, J. P. (1996). Conceptual automaticity in recognition memory: Levels of processing effects on familiarity. *Canadian Journal of Experimental Psychology*, **50**, 123–138.
- Toth, J. P. (2000). Nonconscious forms of human memory. In E. Tulving & F. I. M. Craik (Eds.), *Oxford handbook of memory* (pp. 245–261). New York: Oxford Univ. Press.
- Toth, J. P., Reingold, E. M., & Jacoby, L. L. (1994). Toward a redefinition of implicit memory: Process dissociations following elaborative processing and self-generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 290–303.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, **4**, 207–232.
- Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **19**, 1235–1253.
- Whittlesea, B. W. A. (1997). Production, evaluation, and preservation of experiences: Constructive processing in remembering and performance tasks. In D. Medin (Ed.), *Psychology of Learning and Motivation*, **37**, 211–264.
- Whittlesea, B. W. A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? A discrepancy-attribution account of feelings of familiarity. *Acta Psychologica*, **98**, 141–165.
- Whittlesea, B. W. A., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **26**, 547–565.
- Witherspoon, D., & Allan, L. G. (1985). The effects of a prior presentation on temporal judgments in a perceptual identification task. *Memory & Cognition*, **13**, 101–111.
- Weldon, M. S., Roediger, H. L., Beitel, D. A., & Johnston, T. R. (1995). Perceptual and conceptual processes in implicit and explicit tests with picture fragment and word fragment cues. *Journal of Memory and Language*, **34**, 268–285.
- Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluation. *Psychological Bulletin*, **116**, 117–142.
- Wilson, T. D., Houston, C. E., Brekke, N., & Etling, K. M. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, **125**, 387–402.
- Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and response speed. *Canadian Journal of Experimental Psychology*, **48**, 516–534.

(Received March 6, 2001)

(Published online January 10, 2002)