

Conflict and Criterion Setting in Recognition Memory

Tim Curran and Casey DeBuse
University of Colorado

P. Andrew Leynes
College of New Jersey

Recognition memory requires both retrieval processes and control processes such as criterion setting. Decision criteria were manipulated by offering different payoffs for correct “old” versus “new” responses. Criterion setting influenced the following late-occurring (1,000+ ms), conflict-sensitive event-related brain potential (ERP) components: the stimulus-locked late posterior negativity (LPN) and the response-locked error-related negativity (ERN). The LPN–ERN were most negative to hits under conservative payoff conditions involving conflict between the correct old response and the payoff for new responses. This same conservative–hit condition was most frequently associated with response reversals when fast initial judgments were followed by slower judgments. Postresponse ERP activity may index conflict-sensitive processes underlying postretrieval cognitive control mechanisms involved with assessing responses to current items and updating response criteria on later trials.

Keywords: memory, ERP, ERN, response bias

Human memory depends on the interplay between memory processes and control processes, as exemplified by the signal detection theory perspective of recognition memory. Old (studied) and new (nonstudied) items differ along a continuum of memorability (e.g., retrieval strength or familiarity). Insofar as the strength of old and new items overlap, strength alone cannot serve as the basis for recognition judgments, so a criterion-setting process is needed to determine the level of memory strength required for an affirmative judgment. Although one important goal is to understand the underlying memory processes, it is equally important to understand how response bias (or how memory is evaluated) affects remembering. For example, response bias may differ for special populations even when memory processes are equivalent between groups (Snodgrass & Corwin, 1988).

Event-related brain potentials (ERPs) have been used to investigate the brain processes that discriminate old and new items on recognition memory tasks. About 300–800 ms after onset of a test item, ERPs to correctly recognized old items (hits) are more positive than those to correctly rejected new items. These ERP old–new effects are thought to reflect the activity of retrieval processes that might support the behavioral discrimination of old and new items on recognition memory tests. In particular, distinct

ERP old–new effects have been hypothesized to be associated with separate recollection and familiarity processes that are thought to underlie recognition memory (Jacoby, 1991; Malmberg, Holden, & Shiffrin, 2004; Norman & O’Reilly, 2003; Reder et al., 2000; Rugg & Yonelinas, 2003; Yonelinas, 2002). Several studies have distinguished between midfrontal ERP old–new difference peaking around 400 ms (*FN400 old–new effect*), and a parietal old–new difference peaking around 600 ms (*parietal old–new effect*). It has been suggested that the FN400 old–new effect is related to familiarity, whereas the parietal old–new effect is related to recollection (Curran, 2000; Curran & Cleary, 2003; Nessler, Mecklinger, & Penney, 2001, 2005; Rugg et al., 1998; Rugg & Yonelinas, 2003), although others have disagreed (Olichney et al., 2000; Yovel & Paller, 2004).

Later (~1,000+ ms) memory-related ERP effects have been hypothesized to be related to memory control processes. First, late, frontal, and often right-lateralized ERP old–new effects are thought to reflect the activity of postretrieval evaluation processes (Allan, Wilding, & Rugg, 1998; Curran & Friedman, 2003; Curran, Schacter, Johnson, & Spinks, 2001; M. K. Johnson, Kounios, & Nolde, 1996; Leynes, Bink, Marsh, Allen, & May, 2003; Ranganath & Paller, 2000; Wilding, 1999; Wilding & Rugg, 1997a, 1997b). Second, a late posterior negativity (LPN) effect typically manifests itself as a greater negativity to old than new items. On the basis of a review of the literature, Johansson and Mecklinger (2003) concluded that the LPN is typically observed under one of two conditions: (a) when response conflict exists because of inconsistent sources of information (also see, R. Johnson, Barnhardt, & Zhu, 2004, 2005) or (b) when the memory task requires some binding of items to contextual information as in source judgments (also see, Friedman, Cycowicz, & Bersick, 2005; Wilding & Rugg, 1997b). In the former situation, the stimulus-locked LPN may reflect the activity of the same processes underlying the response-locked error related negativity (ERN). Whereas stimulus-locked ERPs are computed by averaging the electroencephalogram (EEG) from individual trials that are time locked to stimulus onset, response-locked ERPs are computed by averaging the EEG from

—Tim Curran and Casey DeBuse, Department of Psychology, University of Colorado; P. Andrew Leynes, Department of Psychology, College of New Jersey.

This research was supported by National Institutes of Mental Health Grant MH64812 and by the James S. McDonnell Foundation. We thank Dan Blum, Sophie Boddington, John Capps, Debra Chase, Chris D’Lauro, Danielle Germain, Tatsuko GoHollo, Jane Hancock, Gabriel Matthews, Rachel Newton, Christel Taylor, Elan Walsky, Brion Woroch, and Brent Young for research assistance, and Lew Harvey, Michael Frank, Nicole Speer, Rachel Diana, Heekyeong Park, Lynne Reder, Caren Rotello, and Trish Van Zandt for helpful comments.

Correspondence concerning this article should be addressed to Tim Curran, Department of Psychology, University of Colorado, 345 UCB, Boulder, CO 80309-0345. E-mail: tim.curran@colorado.edu

individual trials that are time-locked to the subject's response. The ERN has been thought to reflect response conflict in a variety of tasks (Bartholow et al., 2005; Frank, Woroch, & Curran, 2005; van Veen & Carter, 2002; Yeung, Cohen, & Botvinick, 2004), so this is generally consistent with its hypothesized relationship to conflicting memory judgments.

A goal of the present research is to provide more convincing evidence that directly links the LPN-ERN to memory-related response conflict because prior research has provided some weak support for this connection. The ERN is commonly estimated to arise from medial frontal cortical regions including the anterior cingulate (Dehaene, Posner, & Tucker, 1994; Holroyd, Dien, & Coles, 1998; Luu, Tucker, Derryberry, Reed, & Poulsen, 2003), and a recent functional magnetic resonance imagery study of memory has implicated the anterior cingulate in memorial conflict resolution (Bunge, Burrows, & Wagner, 2004). Studies of cognitive control have indicated that the response-locked ERN is larger when response ambiguity exists (Scheffers & Coles, 2000), and some memory studies similarly have found enhanced stimulus-locked LPNs and/or response-locked ERNs under conditions associated with response ambiguity. First, the LPN-ERN is larger for false alarms to semantically similar lures than for hits or for correct rejections, and this has been interpreted to reflect greater response conflict for similar lures that are difficult to unambiguously categorize as old versus new (Johansson & Mecklinger, 2003; Nessler & Mecklinger, 2003). Second, Johnson and colleagues (R. Johnson, Barnhardt, & Zhu, 2003, 2004, 2005) have consistently found larger response-locked ERNs recorded during recognition memory tests in which subjects are instructed to respond deceptively rather than truthfully. These conditions presumably create a conflict between the subject's actual memory assessment and the instruction to respond deceptively.

Just as a competition between truthfulness and deception created conflict in these experiments, it is conceivable that inducements to maintain a liberal or conservative response criterion may sometimes conflict with information retrieved from memory about individual items (e.g., conflict between a conservative response criterion and evidence that a given item is truly old). The present experiments examine this possible source of conflict by arranging payoffs that induce liberal or conservative response biases on different tests of recognition memory.

Response bias has been addressed in only a few ERP studies of recognition memory. Windmann and Kutas (2001) found that 300- to 500-ms frontal ERP differences between items associated with hits versus false alarms (presumably related to the FN400) were larger for neutral words than for negative words. They interpreted this difference to reflect more liberal responding to emotional words, as was observed in their behavioral results and other results (Maratos, Allan, & Rugg, 2000; Windmann & Kruger, 1998). In another study, subjects were split according to their response bias, and 300- to 500-ms frontal ERP old-new effects (correct trials only, as is typical of ERP recognition memory studies) were larger for conservative than liberal subjects (Windmann, Urbach, & Kutas, 2002). Because any of the aforementioned effects could be attributed to factors other than differences in criterion setting, Azimian-Faridani and Wilding (2006) directly manipulated response criteria by instructing subjects to respond liberally in some blocks (i.e., "make a new response only when confident that the test word was new. . .") and conservatively in other blocks. A 300-

to 500-ms midfrontal FN400 old-new effect was observed in both conditions, but overall amplitudes were more positive going in the conservative than liberal condition. Thus, a few have suggested that response bias may influence the FN400 old-new effect, which is hypothesized to be related to familiarity.

We sought to better understand the effects of criterion setting on memory-related ERP effects by manipulating response criteria with a payoff manipulation (Healy & Kubovy, 1978; Macmillan & Creelman, 2005) rather than using confidence (Azimian-Faridani & Wilding, 2006). Furthermore, Azimian-Faridani and Wilding examined only ERPs from 300–900 ms, so we also were interested in assessing the possibility that later ERP effects that have been associated with memory control processes might specifically be influenced by the payoff manipulation. In particular, we were interested in possible conflict arising between instructions to favor old versus new responses versus the actual old-new status of items, as indexed by the LPN-ERN. To accomplish these goals, we gave the subjects points for each correct response, and they were told that the participant with the highest point total would win \$25. In liberal blocks, subjects were awarded 5 points for each hit and 1 point for each correct rejection. In conservative blocks, subjects were awarded 1 point for each hit and 5 points for each correct rejection.

Experiment 1

Method

Subjects

Forty-seven University of Colorado students participated in the experiment for payment or course credit. Data were discarded from 4 subjects who did not adhere to instructions and failed to set their response criteria higher for conservative than for liberal blocks (as detailed below). Other subjects were excluded for concussion history ($n = 2$), excessive blinking, computer malfunction, or less than 20 acceptable trials per condition. Of the 38 subjects included in analyses, 10 were women, and all were right handed.

Design

Subjects completed four study-test blocks in a 2 Old-New Effects \times 2 Response Bias (conservative, liberal) within-subjects design. Response bias was varied across alternating odd-even test blocks with order counterbalanced across subjects.

Stimuli

Common English words were presented in the center of a flat panel computer monitor. The 320 words were divided into four sets of 80 that were rotated across the 4 Old-New \times Conservative-Liberal conditions across subjects. Each set was roughly equated for length ($M = 5.41$ letters, $SD = 1.05$, range = 4–7) and word frequency ($M = 15.88$, $SD = 19.16$, range = 1–99; Kucera & Francis, 1967). Other words with similar characteristics were used for practice and buffer items.

Procedure

Each session lasted about 1.5–2.0 hr. Subjects completed a consent form, after which they were set up with the geodesic sensor net and completed a brief practice block before the main experiment began. Sensor net setup typically takes about 20–30 min and includes placing the net on the

Table 1
Experiment 1 and 2 Behavioral Results

Experiment	Condition	Hits	False alarms	d'	c	RT new	RT old
1	Conservative	0.73	0.17	1.69	0.17	2,474	2,449
	Liberal	0.87	0.34	1.67	-0.38	2,476	2,446
2	Conservative	0.72	0.14	1.80	0.25	1,159	1,122
	Liberal	0.84	0.29	1.71	-0.25	1,108	980

Note. RT = reaction time.

subject's head, positioning sensors, and adjusting and/or rewetting sensors to reach desired impedance levels of less than 40 k Ω .

Each study list contained 40 experimental words surrounded by two word-primacy and recency buffers that were not later tested. Each study trial included a 200-ms fixation sign (+), followed by the word for 1 s, and a 100-ms blank screen between trials. Subjects were asked to count the syllables in each word and press the corresponding response key (1, 2, 3, or 4) before the next trial began. We chose this shallow encoding task so that accuracy was low enough for responses to be strongly influenced by the payoffs. A 2-min rest break intervened between each study and test list.

Each 80-word test list contained the 40 studied (old) words randomly intermixed with 40 nonstudied (new) words. Subjects received a self-paced rest break after every 20 trials to rest their eyes. Test trials began with a fixation (+) that randomly varied between 500–1,000 ms, followed by a word for 2 s, and then a question mark until the response. Subjects were asked to delay their response until the question mark appeared. Responses were given with the first finger of each hand, with old–new to left–right assignment counterbalanced across subjects.

Throughout each test block, possible point rewards were displayed on the bottom of the screen. Subjects were not informed of the point values until after the study list. For conservative blocks, subjects were rewarded 5 points for each correct “new” response and 1 point for each correct “old” response. Points were reversed for liberal blocks. After each test block, a subject's current point total was displayed. Subjects were told that the participant with the highest final score would be rewarded \$25 after the experiment was completed, and \$25 was accordingly awarded to the highest scoring subject.

EEG–ERP Recording and Analysis

During the recognition memory task, scalp voltages were collected with a 128-channel geodesic sensor net (Tucker, 1993) connected to an AC-coupled, 128-channel, high-input impedance amplifier (200 M Ω , Net Amps, Electrical Geodesics, Eugene, OR). Amplified analog voltages (0.1- to 100.0-Hz band-pass) were digitized at 250 Hz. Individual sensors were adjusted until impedances were less than 50 k Ω . The EEG was digitally low-pass filtered at 40 Hz. Trials were discarded from analyses if they contained incorrect responses, eye movements (eye channel amplitudes over 70 μ V), or more than 20% of channels were bad (average amplitude over 100 μ V or transit amplitude over 50 μ V). Individual bad channels were replaced on a trial-by-trial basis with a spherical spline algorithm (Srinivasan, Nunez, Silberstein, Tucker, & Cadusch, 1996). EEG was measured with respect to a vertex reference (Cz), but an average-reference transformation was used to minimize the effects of reference-site activity and accurately estimate the scalp topography of the measured electrical fields (Dien, 1998; Picton, Lins, & Scherg, 1995). The average reference was corrected for the polar average reference effect (Jung-höfer, Elbert, Tucker, & Braun, 1999).

ERPs were obtained by stimulus-locked averaging of the EEG recorded in each condition. ERPs were baseline-corrected with respect to a 100-ms prestimulus recording interval. The number of trials per subject per condition used to calculate ERPs were as follows: conservative–new ($M = 54.13$, $SD = 11.16$, range = 29–75), conservative–old ($M = 47.87$, $SD =$

11.83, range = 24–72), liberal–new ($M = 42.63$, $SD = 12.21$, range = 22–65), and liberal–old ($M = 57.95$, $SD = 11.06$, range = 34–78).

Results

Behavioral Results

Behavioral results are summarized in Table 1. Both hit rates (H), $t(37) = 7.14$, $SE = .02$, $p < .001$, and false alarm rates (F), $t(37) = 7.63$, $SE = .02$, $p < .001$, were higher for liberal than for conservative tests. Response bias was estimated with a criterion location measure, defined as $c = -.5[z(H) + z(F)]$ (Macmillan & Creelman, 2005). Negative values indicate a liberal response bias, whereas positive values indicate a conservative bias, and the present bias conditions differed accordingly, $t(37) = 8.52$, $SE = .06$, $p < .001$. Discrimination, estimated as $d' = z(H) - z(F)$, did not differ between conditions, $t(37) = .26$, $SE = .07$, $p > .10$.

Despite the fact that subjects were instructed to withhold their response until after the 2-s test-word display, reaction times (RT) differed between old and new items. RT on correct trials was entered into a Bias (conservative, liberal) \times Old–New repeated measures analysis of variance (ANOVA). Subjects were faster to accurately respond to old items than to new items, $F(1, 37) = 9.36$, $MSE = 6,357$, $p < .01$. No other RT effects approached significance.

ERP Results

FN400 old–new effects. The FN400 was analyzed in a priori regions of interest (ROI) established in past research (Curran, 2000, 2004; Curran & Cleary, 2003; Curran, DeBuse, Woroch, & Hirshman, 2006; Curran & Dien, 2003; Curran & Friedman, 2004; Curran, Tanaka, & Weiskopf, 2002).¹ The ROIs were the left and right anterior, superior regions shown in Figure 1 (electrode layout) and Figure 2 (ERPs averaged within each region). Mean amplitude from 300–500 ms was computed by averaging the channels within each region for each condition and/or subject (see Figure 3 left). These means were entered into a Bias (conservative, liberal) \times Old–New \times Hemisphere ANOVA. For this and all subsequent ANOVAs, effects involving merely spatial factors (not interacting with bias or old–new differences) are not reported. Results indicated that anterior, superior amplitudes were more negative for new than old words, $F(1, 37) = 8.44$, $MSE = 2.2$, $p < .01$, and no criterion effects approached significance ($F < 1$).

¹ The broader topography of the FN400, parietal, and LPN effects were also examined across all eight regions shown in Figures 1 and 2, but these analyses did not qualify the conclusions reached from the focused ROIs.

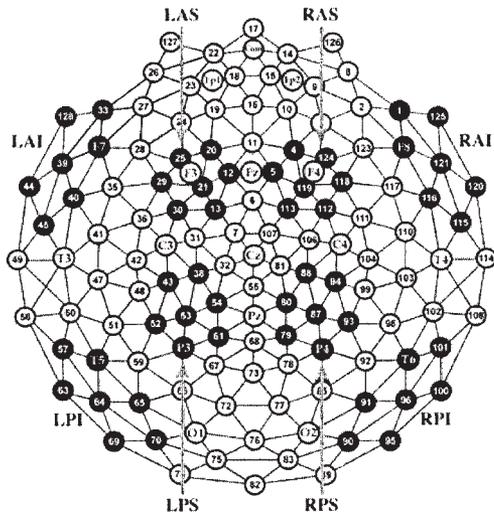


Figure 1. Geodesic sensor net layout. Electrode sites are numbered along with selected 10-10 positions. Black clusters are regions included in the eight-region analyses. L = left; R = right; A = anterior; P = posterior; I = inferior; S = superior. Reprinted from “Effects of Attention and Confidence on the Hypothesized ERP Correlates of Recollection and Familiarity,” by T. Curran, 2004, *Neuropsychologia*, 42, pp. 1088–1106. Copyright 2004 by Elsevier. Reprinted with permission.

Parietal old–new effects. The parietal old–new effects were analyzed from 500–700 ms within left and right posterior, superior ROIs (see Figures 1 and 2). Posterior, superior amplitudes were more positive for old than new words, $F(1, 37) = 8.49$, $MSE = 4.56$, $p < .01$ (Figure 3 middle), and the Criterion \times Old–New \times Hemisphere interaction approached significance, such that the left-hemisphere old–new differences tended to be larger in the conservative blocks, $F(1, 37) = 3.52$, $MSE = 0.23$, $p = .07$ (compare left and right posterior, superior ERPs in Figure 2).

LPN. The LPN was analyzed across the same parietal (LPS, RPS) ROIs because it is typically maximal over these regions. Because of its broad time course, we added a temporal factor to an ANOVA that was otherwise the same as those described previously: Time (1,000–1,500 ms, 1,500–2,000 ms) \times Bias (conservative, liberal) \times Old–New \times Hemisphere. LPN amplitude was more negative for conservative than liberal blocks, $F(1, 37) = 5.67$, $MSE = 5.26$, $p < .05$; and was more negative for old than for new words, $F(1, 37) = 9.61$, $MSE = 4.56$, $p < .01$. The old–new differences were greater over the left than the right hemisphere, $F(1, 37) = 13.13$, $MSE = 0.68$, $p < .001$; greater from 1,000–1,500 ms than from 1,500–2,000 ms, $F(1, 37) = 9.48$, $MSE = 0.65$, $p < .01$; and greater for conservative than for liberal conditions, $F(1, 37) = 5.28$, $MSE = 4.41$, $p < .05$ (see Figure 3, right).

Discussion

The payoff manipulation had the desired effect on response criteria without significantly influencing discrimination. Response criteria were higher in the conservative condition than in the liberal condition, but discrimination estimates were statistically equivalent between the two conditions. These conclusions must be tempered by the inaccuracy of single-point measures of discrimination

(d') and response bias (c). When the slope of the z -transformed receiver operating characteristic (z ROC) curve equals one (implying equal variance of signal and noise distributions), d' and c should be accurate (Macmillan & Creelman, 2005). However, z ROC slopes near .80 are typically observed in recognition memory (Glanzer, Kim, Hilford, & Adams, 1999; Ratcliff, Shue, & Gronland, 1992), so our single-point estimates are possibly inaccurate. This issue will be addressed in Experiment 3 by using confidence judgments to construct ROC curves that provide better estimates of discrimination and response bias under similar conditions. As will be detailed later, these results support the present conclusion that the payoffs influenced response bias without affecting discrimination.

Response criteria had negligible influence on the ERP old–new effects prior to 1,000 ms (i.e., the marginally significant Bias \times Old–New \times Hemisphere interaction on the parietal effect that was not replicated in Experiment 2), but the old–new difference was larger for conservative than for liberal response criteria in the 1,000- to 1,500-ms LPN. The late time course of these effects suggests that altering criterion by payoffs more likely produced changes in the postretrieval metacognitive processes rather than influencing the evaluation of familiarity, which would influence ERP activity in the 300- to 800-ms time window. Perhaps the influence of response criteria was largest for the LPN because of the conflict that may arise between a conservative bias and “old” responses. A conservative response criterion biases subjects to respond “new,” which will conflict with “old” responses to items that are accurately remembered as studied items. By this logic, one might expect a similarly enhanced LPN on liberal–new trials, but no such enhancement was observed. The behavioral results provide a hint that subjects might have felt more conflicted during conservative blocks because they were less successful in moving their criterion upward (conservative $c = .17$) than downward (liberal $c = -.38$). This conflict may result from a departure from one’s natural response tendency on the recognition test. More specifically, subjects might have been naturally more likely to adopt a liberal criterion on the test. This point is more thoroughly addressed by Experiment 3.

The plausibility of any explanation based on conflict would be enhanced if it could be shown that in the present experiments that the stimulus-locked LPN was related to the response-locked ERN (Johansson & Mecklinger, 2003). More specifically, the argument that subjects experienced increased conflict in the conservative–old condition would be stronger if the ERN in this condition was also larger. Unfortunately, the delayed response procedure used in the present experiment precludes the derivation of meaningful response-locked ERPs. Thus, the goal of Experiment 2 was to replicate Experiment 1 with a response procedure that would allow for measurement of the response-locked ERN.

Experiment 2

Method

The method of Experiment 2 was the same as Experiment 1, except that Experiment 2 subjects were instructed to respond as soon as they made their decision rather than to wait 2 s for test-word offset.

Thirty-seven University of Colorado students participated in the experiment for payment. Data were discarded from 2 subjects who failed to set their response criteria higher for conservative than for liberal blocks. Other subjects were excluded for excessive blinking ($n = 2$), low number of

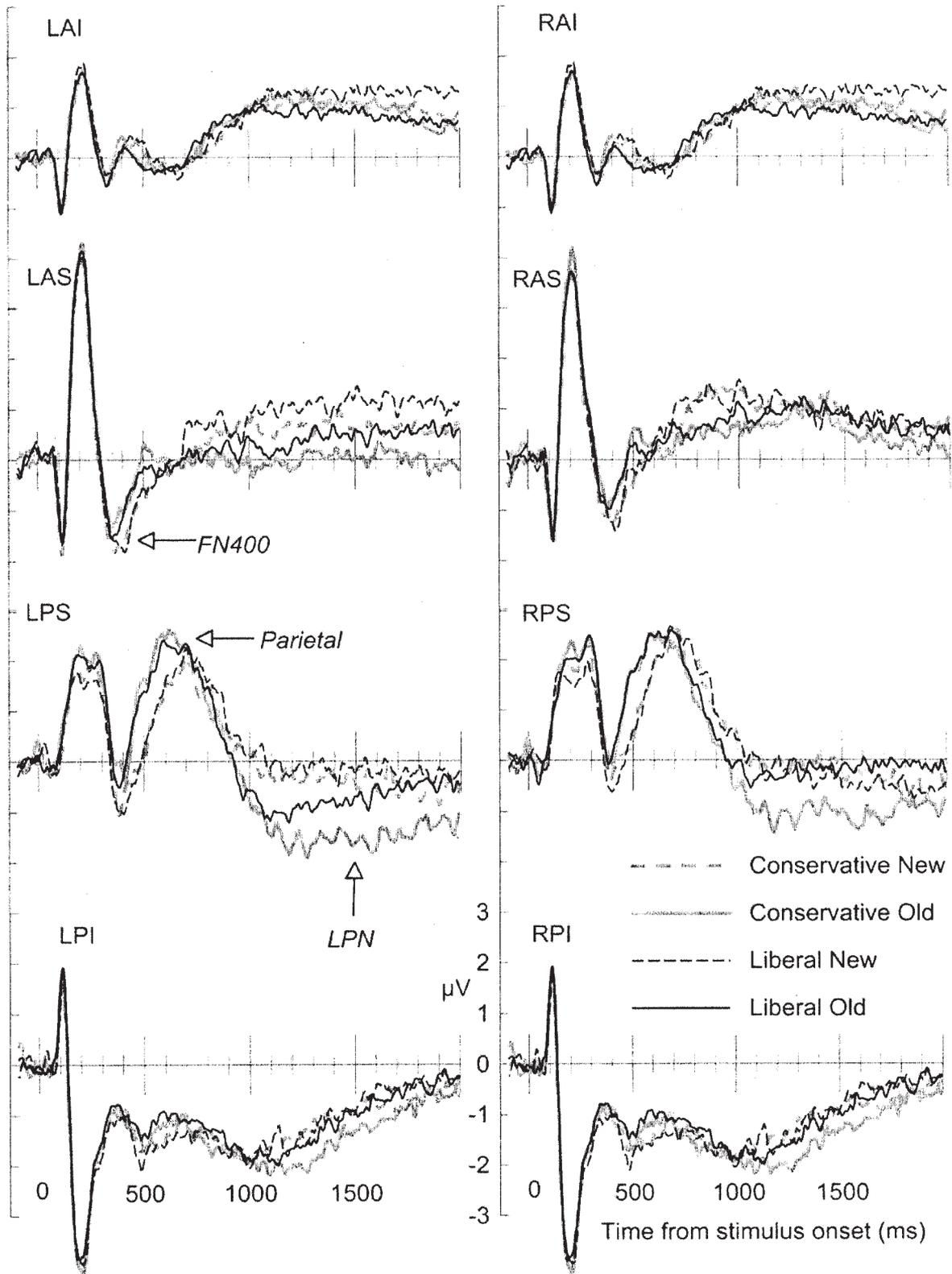


Figure 2. Experiment 1 event-related potentials (ERPs). Grand-averaged ERPs in each of the eight regions depicted in Figure 1. L = left; R = right; A = anterior; P = posterior; I = inferior; S = superior; FN400 = midfrontal ERP old-new difference peaking around 400 ms; LPN = late posterior negativity.

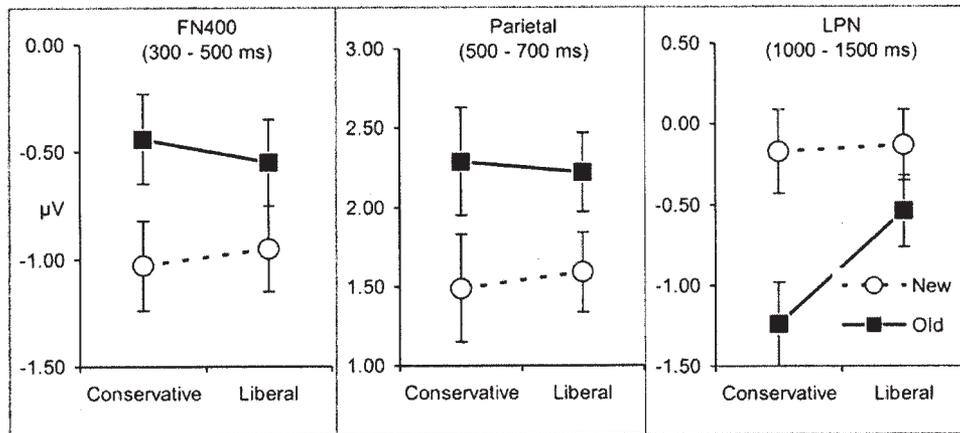


Figure 3. Experiment 1 mean amplitudes. Error bars are the standard error of the old–new difference. Left panel: left and right anterior, superior regions, 300–500 ms. Middle panel: left and right posterior, superior regions, 500–700 ms. Right panel: left and right posterior, superior regions, 1,000–1,500 ms. FN400 = midfrontal event-related potential old–new difference peaking around 400 ms; LPN = late posterior negativity.

artifact-free trials because of low accuracy ($d' = .41$), excessive EEG noise, or hydrocephalus. Of the 32 subjects included in analyses, 11 were women, and all were right handed.

Results

Behavioral Results

Behavioral results are summarized in Table 1. Both hit rates, $t(31) = 7.58$, $SE = .02$, $p < .001$, and false alarm rates, $t(31) = 7.79$, $SE = .02$, $p < .001$, were higher for liberal than for conservative tests. Also, c differed as expected between liberal and conservative conditions, $t(37) = 9.74$, $SE = .05$, $p < .001$, and d' did not differ significantly between conditions, $t(37) = 1.50$, $SE = .06$, $p = .14$.

RT was faster in liberal than in conservative blocks, $F(1, 31) = 5.94$, $MSE = 50,037$, $p < .05$, and faster for old than for new words, $F(1, 31) = 8.28$, $MSE = 26,342$, $p < .01$, and these two factors significantly interacted, $F(1, 31) = 9.72$, $MSE = 6,730$, $p < .01$. The interaction was due to the liberal old condition being significantly faster than each of the other conditions (all $ps < .01$).

ERP Results

FN400 old–new effects. ERP waveforms are shown in Figure 4. Anterior, superior amplitudes were more negative for new than old words, $F(1, 31) = 40.69$, $MSE = 1.26$, $p < .001$, and no criterion effect approached significance (see Figure 5, left). Old–new differences were larger over the left than the right hemisphere, $F(1, 31) = 7.09$, $MSE = 0.16$, $p < .05$.

Parietal old–new effects. Posterior, superior amplitudes were more positive for old than new words, $F(1, 31) = 44.84$, $MSE = 3.12$, $p < .001$, and this difference was larger over the left than the right hemisphere, $F(1, 31) = 6.37$, $MSE = .28$, $p < .05$. Marginally significant bias effects captured a trend toward more positive amplitudes for liberal than for conservative conditions, $F(1, 31) = 3.70$, $MSE = 2.86$, $p = .06$ (see Figure 5, middle). No interactions between bias and old–new differences approached significance.

LPN. Posterior, superior amplitudes were more negative for old than for new words, $F(1, 31) = 4.75$, $MSE = 4.92$, $p < .05$, and were more negative for conservative than for liberal blocks, $F(1, 31) = 5.33$, $MSE = 6.79$, $p < .05$ (see Figure 5, right). The Time \times Bias \times Old–New interaction indicated that old–new differences were greatest for the conservative condition between 1,000 and 1,500 ms, $F(1, 31) = 6.46$, $MSE = 0.37$, $p < .05$. The Time \times Bias \times Hemisphere interaction indicated that the LPN was more negative from 1,000–1,500 ms than from 1,500–2,000 ms over both hemispheres for the conservative blocks, but only over the left hemisphere in the liberal blocks, $F(1, 31) = 7.24$, $MSE = 0.10$, $p < .05$.

ERN. Response-locked ERPs were baseline corrected with respect to a -200 -ms preresponse interval. Following previous studies (Frank et al., 2005; Tucker, Luu, Frishkoff, Quiring, & Poulsen, 2003), the ERN was measured within a midcentral group of channels surrounding Fz (Channels 5, 6, 7, 12, 13, 107, and 113; see Figure 1 for locations and Figure 6 for ERP waveforms). Peak amplitude was measured within a 0- to 120-ms postresponse interval. An Old–New \times Bias ANOVA revealed no significant latency differences between conditions, so mean amplitude was quantified within a 40- to 80-ms postresponse window ($M \pm 2SD$, rounded to nearest samples). The Old–New \times Bias interaction was highly significant, $F(1, 31) = 18.40$, $MSE = 0.79$, $p < .001$. The ERN was more negative for the conservative old condition compared with each of the others (all $ps < .001$, see Figure 7).

Discussion

As in Experiment 1, payoffs influenced response criteria without significantly affecting discrimination. Although a marginally significant trend in Experiment 1 suggested that the parietal old–new effect may be larger for conservative than for liberal blocks, that trend was not replicated in Experiment 2, so we consider it to be unreliable. Neither experiment replicated a previous finding that the FN400 was more negative in a liberal

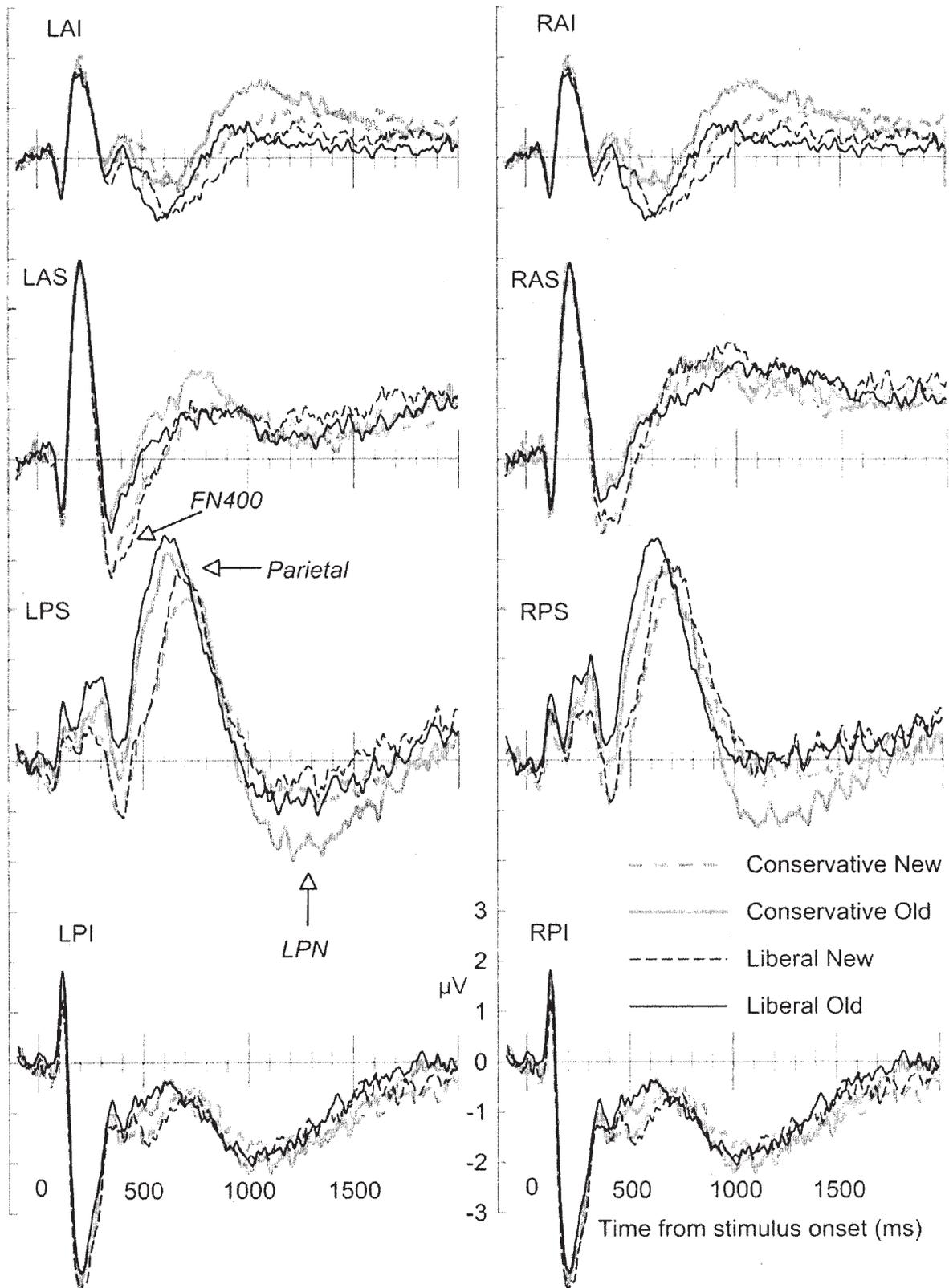


Figure 4. Experiment 2 event-related potentials (ERPs). Grand-averaged ERPs in each of the eight regions depicted in Figure 1. L = left; R = right; A = anterior; P = posterior; I = inferior; S = superior; FN400 = midfrontal ERP old–new difference peaking around 400 ms; LPN = late posterior negativity.

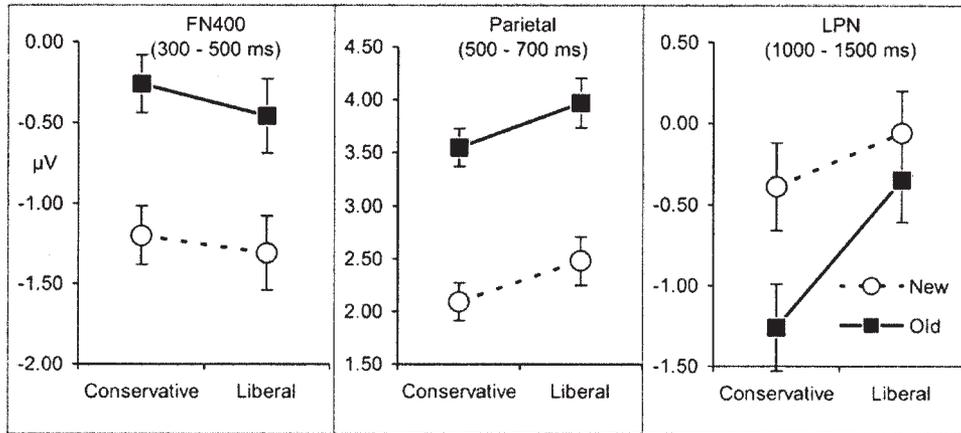


Figure 5. Experiment 2 mean amplitudes. Error bars are the standard error of the old–new difference. Left panel: left and right anterior, superior regions, 300–500 ms. Middle panel: left and right posterior, superior regions, 500–700 ms. Right panel: left and right posterior, superior regions, 1,000–1,500 ms. FN400 = midfrontal event-related potential old–new difference peaking around 400 ms; LPN = late posterior negativity.

than in a conservative condition (Azimian-Faridani & Wilding, 2006). Compared with the present experiments, Azimian-Faridani and Wilding used a more semantic encoding task (animacy judgments), had overall higher accuracy rates ($d' > 2$), and their criterion estimates differed more between liberal and conservative conditions. Furthermore, there may be important differences between manipulating criterion by encouraging high confidence responses to old (liberal) or new (conservative) items (Azimian-Faridani & Wilding, 2006) versus payoffs as in the present experiments, as will be considered further in the General Discussion.

The LPN results replicated those of Experiment 1, with the negativity being largest in the conservative–old condition. Furthermore, a strikingly similar pattern was observed for the ERN. These results suggest that participants experienced greater conflict with conservative than with liberal response criteria. Unlike Experiment 1 in which criterion movement was asymmetric, it was perfectly symmetric in Experiment 2 ($c = .25$ vs. $-.25$), so asymmetric response biases may be unlikely to account for the LPN–ERN results. However, RT to old items was slower for

conservative than for liberal trials, which may indicate greater conflict or indecision during conservative blocks. These RT effects could not be observed in Experiment 1 because the subjects' responses were delayed. To address the idea that subjects might be more naturally inclined to respond liberally under the present conditions, in Experiment 3 we included a first block with equal payoffs for hits and correct rejections to induce a neutral response criterion.

The primary goal of Experiment 3 was to provide direct evidence that subjects are experiencing more conflict in the conservative–old condition. To this point, this interpretation has

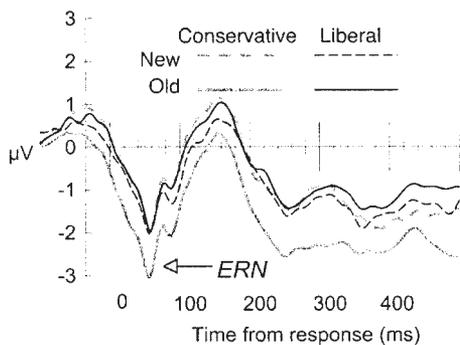


Figure 6. Experiment 2 error-related negativity (ERN). Grand average event-related potentials obtained by averaging across Channels 5, 6, 7, 12, 13, 107, and 113 (see Figure 1 for locations).

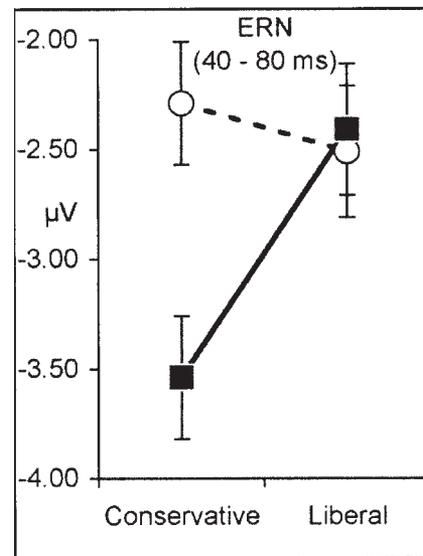


Figure 7. Experiment 2 error-related negativity (ERN) mean amplitudes. Error bars are the standard error of the old–new difference. Means were obtained by averaging across Channels 5, 6, 7, 12, 13, 107, and 113 (see Figure 1 for locations). Squares = old condition; circles = new condition.

rested on speculation informed by past research relating the ERN to conflict (e.g., Bartholow et al., 2005; Frank et al., 2005; van Veen & Carter, 2002; Yeung et al., 2004). Here, we sought behavioral evidence of conflict under the same conditions producing the ERP effects in Experiments 1 and 2 by using a dual-response procedure that allows subjects to reverse their initial responses (following Van Zandt & Maldonado-Molina, 2004).

Van Zandt and Maldonado-Molina (2004) provided compelling evidence that subjects can experience response conflict during recognition memory judgments that results from a continuation of information processing after a response has been made. Their subjects made two responses on each trial (with foot pedals and manually). Critically, subjects sometimes reversed their responses between their first and second responses when instructed to give fast (500–700 ms) first responses. To test the hypothesis that the LPN–ERN results in Experiments 1 and 2 reflect response conflict, in Experiment 3 we used a modified version of Van Zandt and Maldonado-Molina’s procedure. If the larger LPN–ERN in the conservative conditions reflects greater response conflict, then response reversals should be greatest for this condition. Following Van Zandt & Maldonado-Molina’s Experiment 1, confidence judgments were required for the second response.²

Experiment 3

Method

Subjects

Twenty-five University of Colorado students participated in the experiment for payment. Data from 5 subjects were discarded because (a) they failed to set their response criteria higher for conservative than liberal blocks ($n = 2$), (b) exceedingly fast RT on second responses suggested that they preprogrammed the first and second responses together ($n = 2$; $M_1 = 90$ ms, $M_2 = 205$ ms), or (c) the computer crashed during the experiment ($n = 1$). Of the 20 subjects included in analyses, 14 were women.

Design

Subjects completed five study-test blocks in a 2 Old–New \times 3 Response Bias (conservative, liberal, neutral) within-subjects design. One neutral block was followed by two conservative and two liberal blocks. The neutral block was always first because it was primarily intended to measure each subject’s natural response tendency before experiencing the asymmetric payoffs. Thus, because of this order confound, the neutral condition is not directly comparable with the others. Order of the liberal and conservative blocks (2–5) was counterbalanced in an alternating fashion as in the previous experiments. Within each test, half of the words were old (studied), and half were new (not studied).

Stimuli

Another set of 80 words with similar characteristics was added to the four sets of 80 described in the Experiment 1 *Method* section. The five sets were roughly equated for length ($M = 5.41$ letters, $SD = 1.06$, range = 4–7) and word frequency ($M = 16.07$, $SD = 19.53$, range = 1–99; Kucera & Francis, 1967). These five sets were completely counterbalanced through the three response bias and two old–new conditions. Other words with similar characteristics were used for practice and buffer items.

Procedure

Each session lasted 45–60 min. Subjects completed a consent form, followed by three practice blocks and the main experiment. Study blocks

were identical to those in the first two experiments. Test trials were modified such that subjects responded twice on each trial. Subjects situated their hands on a standard keyboard, including the numeric keypad to right, such that their thumbs were positioned over the space bar (left hand) and *O* (right hand) keys, while the first three fingers of each hand were positioned over the *Z*, *X*, and *C*, (left) and the *I*, *2*, and *3* (right) keys. On each trial, subjects made a fast (700-ms deadline) old–new response with their thumbs followed by a confidence judgment with their fingers.³

The first and second practice blocks were intended to acquaint the subject with the key assignments and speeded response procedure. The first practice block consisted of 20 trials in which a 500- to 1,000-ms fixation sign (+) was followed by either the word “Left” or “Right.” Subjects were instructed to respond with a right or left thumb press within 700 ms of seeing the corresponding word. After responding, the subject was given feedback on RT and accuracy. The second practice block included 40 similar trials except that the words “Left” and “Right” were replaced with “Studied” and “Nonstudied.” The third practice block included a 10-word study list followed by a 40-word test list. Immediately after the third practice block, as well as every block of the experiment thereafter, the subject received feedback about the percentage of first responses made within 700 ms, the mean first response RT, and the mean accuracy of his/her first responses for that block.

For actual test trials, the subject’s first old–new response was followed by a confidence rating. Immediately after the first response, a 6-point confidence interval appeared onscreen, along with the message “Too Slow,” for 500 ms if their initial judgment was after the 700-ms deadline. The terms “Surely Studied,” “Likely Studied,” “Maybe Studied,” “Maybe Nonstudied,” “Likely Nonstudied,” and “Surely Nonstudied” appeared in positions corresponding to the response keys. Subjects made this non-speeded confidence response using the *Z*, *X*, and *C* keys and the *I*, *2*, and *3* keys on the number pad with their first three fingers of both hands. The assignment of studied and nonstudied confidence choices to left–right hand keys was counterbalanced across subjects, such that they matched the initial studied–nonstudied thumb assignment.

Throughout each test block, possible point rewards were displayed on the screen. For the first neutral block, subjects were rewarded 1 point for each correct “Nonstudied” response and 1 point for each correct “Studied” response. For conservative blocks, subjects were rewarded 5 points for each correct “Nonstudied” response and 1 point for each correct “Studied” response. Point rewards were reversed for liberal blocks. Points applied to both correct initial studied–nonstudied judgments made within 700 ms, as well as correct studied–nonstudied confidence judgments, regardless of confidence level. In all blocks, 1 point was deducted for initial studied–nonstudied judgments made after 700 ms, regardless of accuracy. After each test block, a subject’s current point total was displayed. As in previous experiments, the subject with the highest point total was awarded \$25.

² We decided to not record EEG during Experiment 3 for several reasons. First, the fast (<700 ms) first responses could potentially contaminate the recording of stimulus-locked ERPs occurring within this timeframe. Second, pilot testing indicated that we would not have enough observations to compute separate ERPs for trials with and without reversals of the first response. This problem would be exacerbated by separating trials according to confidence. Thus, we opted to run a parallel behavioral study to examine subject natural response biases and reversal rates under conditions that were analogous to the first two EEG experiments.

³ Van Zandt and Maldonado-Molina (2004) required subjects to respond between 500 and 700 ms. Our pilot testing suggested that the lower bound needlessly complicated the task for subjects and overly constrained the number of acceptable responses, so we adopted the procedure of merely requiring responses under 700 ms.

Results

Analyses included only trials in which the first response was faster than 900 ms, even though subjects were given a target speed of 700 ms (following Van Zandt & Maldonado-Molina, 2004). This included 95% of neutral trials and 99% of liberal and conservative trials, so the majority of trials were retained. We computed d' and c estimates for first responses, as in previous experiments. Constructing ROC curves using confidence ratings from the second response created more accurate estimates of discrimination (Az) and response bias (c_a ; Macmillan & Creelman, 2005). Lewis Harvey's RscorePlus program, using maximum likelihood estimates, was used to calculate ROC curves and statistics for each subject (Harvey, 2005).

Response Bias

The primary purpose of the first neutral block was to assess subjects' natural response biases under the present conditions, and results revealed a liberal bias for both first and second responses. As seen in Table 2, subjects adopted a liberal (significantly below zero) response bias for their first response, $t(19) = -4.55$, $SE = .06$, $p < .001$. For the second response, Figure 8 (right panel) shows that three of the five neutral criterion location estimates (c_a) were below zero, and the mean of these five locations ($c_a = -.27$) was significantly below zero, $t(19) = -4.34$, $SE = .06$, $p < .001$. Because the neutral block was always first, it is not directly comparable with the following liberal and conservative blocks that were counterbalanced for order. Therefore, inferential statistics below will not include the neutral block when comparing conditions, but the means are presented in the tables and figures for completeness.

For both the first and second responses, response bias differences were consistent with the payoff manipulation. For the first response (Table 2), response bias was lower in the liberal than conservative block, $t(19) = 7.06$, $SE = .13$, $p < .001$. Response bias on the second response was analyzed in a Condition (liberal, conservative) \times Confidence Level (c1, c2, c3, c4, c5) repeated measures ANOVA with c_a as the dependent measure (Figure 8B), using the Geisser–Greenhouse correction for sphericity violations, which is only necessary for effects with more than one degree of freedom. Criterion estimates increased from conservative to liberal conditions, $F(1, 19) = 21.20$, $MSE = .26$, $p < .001$, and increased with confidence level, $F(4, 76) = 80.30$, $MSE = .45$, $p < .0001$. These two factors interacted such that condition differences were larger at the three lower confidence levels, $F(4, 76) = 4.38$, $MSE = .03$, $p < .05$ (see Figure 8B).

Criterion estimates differed between liberal and conservative blocks more for first than second responses. Directly comparing response bias estimates between the first and second responses is

Table 2
Experiment 3 Accuracy on First Response

Condition	Hits	False alarms	d'	c
Conservative	0.52	0.33	0.53	0.20
Liberal	0.82	0.63	0.63	-0.69
Neutral	0.75	0.47	0.75	-0.30

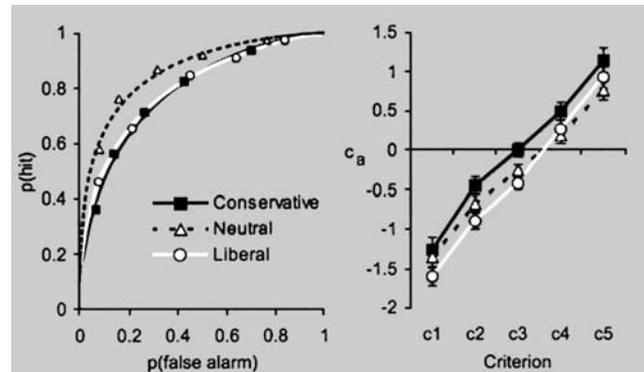


Figure 8. Receiver operating characteristic (ROC) results from the second responses in Experiment 1. Left panel: ROC curves with Az estimates adjacent to the legend. Right panel: Criterion estimates at each level of confidence.

complicated by the different measures used for yes–no judgments (c) versus confidence ratings (c_a). When the slope of the zROC curve equals 1.00 (implying equal variance of signal and noise distributions), c should be equivalent to the middle c_a criterion (Macmillan & Creelman, 2005). Slopes near .80 are typically observed in recognition memory (Glanzer et al., 1999; Ratcliff et al., 1992), and consistent values were observed here (conservative = .86, liberal = .80), so c and c_a are not necessarily equivalent in the present experiment. Thus, we decided to convert the confidence ratings into binary old–new judgments, so that single-point c estimates could be computed for both first and second response. The accuracy of the second-response c estimates was evaluated by comparing them with the more precise middle c_a values with a Bias (conservative, liberal) \times Measure (c , c_a) ANOVA. The c versus c_a differences did not approach significance ($F < 1$), so we felt confident in the appropriateness of using single-point c estimates to compare bias between the first and second responses. Response bias (c) differed between conservative and liberal blocks, $F(1, 19) = 49.73$, $MSE = .18$, $p < .001$, and this difference was larger for the first than second responses, $F(1, 19) = 15.83$, $MSE = .06$, $p < .001$. Thus, the payoffs influenced first responses more than second responses, as would be expected if fast responses are based on less information retrieved from memory.

Discrimination

Discrimination did not differ between liberal and conservative blocks. For the first response (Table 2), the small d' difference between liberal and conservative blocks was not significant, $t(19) = 1.38$, $SE = .07$. The area under the ROC curve (Az) was used to estimate discrimination from the second responses, which did not differ between conservative ($M = .81$) and liberal ($M = .80$) blocks, $t(19) = .67$, $SE = .01$ (see Figure 8A).

RT

RT on correct trials was analyzed for both responses (see Table 3). For the first responses, a Conservative–Liberal \times Old–New interaction indicated that subjects were faster when responding congruently with the payoffs, $F(1, 19) = 17.91$, $MSE = 2101$, $p <$

.001. In conservative blocks, subjects were faster for “new” than for “old” responses, whereas the opposite was observed in liberal blocks. The RT pattern for second responses (measured from onset of confidence rating screen) was similar to that observed in Experiment 2. The Bias \times Old–New interaction, $F(1, 19) = 13.30$, $MSE = 14,570$, $p < .01$, indicated that the liberal–old condition was faster than each of the other conditions (all $ps < .01$).

Response Reversals

The dependent measure of primary interest was the proportion of second responses that were reversals of the first response, as shown for each possible first response type in Table 4. That is, cases in which subjects originally responded “new” and then switched to “old” or vice versa. For example, the top-right cell shows that 48% of first-response misses (incorrect old trials) in the conservative condition were reversed to hits on the second response. Reversals were more frequent when the first responses were incorrect, $F(1, 19) = 79.31$, $MSE = .05$, $p < .001$, and were more frequent for conservative than for liberal blocks, $F(1, 19) = 10.55$, $MSE = .01$, $p < .01$. These factors interacted such that bias differences were evident only when first responses were errors, $F(1, 19) = 6.42$, $MSE = .01$, $p < .05$. Like the ERN results, the reversal rate following incorrect first responses was numerically highest in the conservative–old condition, but the Bias \times Old–New interaction was not significant, $F(1, 19) = 2.24$, $MSE = .01$. Paired comparisons indicated that the conservative–old condition was associated with a significantly higher reversal rate than each of the other conditions: conservative–new, $F(1, 19) = 6.04$, $p < .05$; liberal–new, $F(1, 19) = 13.75$, $p < .01$; and liberal–old, $F(1, 19) = 11.33$, $p < .01$, all $MSEs = .01$. The same qualitative pattern emerges when only incorrect first responses are considered.

A potential problem with the previous analysis concerns the fact that reversal proportions were conditionalized on the accuracy of the first response, and accuracy rates varied substantially between conditions. Thus, the number of observations differed between conditions. To address this issue, we randomly sampled an approximately equal number of trials for each subject per condition, with sample size determined by the condition with the lowest number of trials per subject per condition. This was done separately for trials associated with incorrect ($M = 12$ trials per subject per condition, $SD = 6.47$) versus correct ($M = 26$ trials per subject per condition, $SD = 12.43$) first responses because accurate trials were much more frequent. Results were similar to those including all observations (see Table 5, with the neutral conditions omitted to remove a further constraint on equating observations). A two-tailed t test focusing on trials with incorrect first responses con-

Table 3
Experiment 3 Reaction Time (RT)

Response	Condition	RT new	RT old
First	Conservative	529	553
	Liberal	568	506
	Neutral	624	582
Second	Conservative	878	878
	Liberal	909	713
	Neutral	1,380	1,016

Table 4
Experiment 3 Response Reversals Including All Trials

Condition	Correct first response		Incorrect first response	
	New	Old	New	Old
Conservative	0.07	0.07	0.37	0.48
Liberal	0.06	0.05	0.33	0.35
Neutral	0.06	0.02	0.42	0.53

firmed that these reversal rates were higher for conservative–old than each other condition: conservative–new ($p < .02$), liberal–new ($p < .02$), and liberal–old ($p < .04$). The pairwise comparisons were confirmed with permutation tests (following Van Zandt & Maldonado-Molina, 2004). The reversal rates in the conservative–old condition were higher than the liberal–old ($p = .03$ for all trials, $p = .05$ for frequency-matched analysis) and the liberal–new conditions (both $ps = .02$), and were marginally higher than the conservative–new condition ($p = .055$ for all trials, $p = .05$ for frequency-matched analysis).

Finally, Pearson correlations did not support the hypothesis that higher reversal rates in the conservative–old condition were attributable to subjects’ natural tendency to respond liberally in the neutral block. Although trending in the predicted direction, correlations were not significant for either first ($r = -.17$) or second ($r = -.16$) responses.⁴ Examination of the complete correlation matrix suggests that the only reliable correlations involved reversals of initially correct “new” responses in the neutral condition. Significant negative correlations involving both first and second responses (r range = $-.36$ to $-.67$) indicated that subjects with liberal criterion settings across all blocks were more likely to switch from a correct “new” response to an incorrect “old” response in the neutral block.

Discussion

ROC analyses confirmed that the payoff manipulation had the desired effects on bias without influencing discrimination, thereby confirming the single-point d' and c estimates obtained in the first two experiments despite the often questionable assumptions of these measures (Glanzer et al., 1999; Ratcliff et al., 1992).

When subjects were given neutral payoffs, significantly negative criterion estimates suggested that people were naturally biased to respond liberally under the present conditions. When subjects were asked to give a fast initial response followed by a slow second response, the second responses were more accurate. Criterion estimates differed between liberal and conservative blocks more for first than for second responses, but criterion estimates were more positive in conservative than in liberal blocks in both cases. Thus, subjects were more likely to be influenced by the payoffs when their accuracy was time limited. Accordingly, their initial responses were faster when responding consistently than inconsistently with the payoffs. Subjects were more accurate on

⁴ As suggested by a reviewer, we also attempted to correlate response bias with LPN and ERN amplitudes in Experiments 1 and 2, but these correlations were also not significant.

Table 5
*Experiment 3 Response Reversals With Matched
 Trial Frequencies*

Condition	Correct first response		Incorrect first response	
	New	Old	New	Old
Conservative	0.07	0.06	0.33	0.48
Liberal	0.07	0.04	0.29	0.35

their second responses because initial errors were often corrected when subjects reversed their responses. This error correction occurred most often for the conservative–old condition, as predicted from the previous ERP experiments in which that condition was associated with the most negative LPN–ERN. These results are consistent with the hypothesis that the LPN–ERN indexes response conflict observed under conditions in which subjects are likely to reverse their initial memory judgments. However, the nonsignificant correlation between subjects' response bias and reversal rates did not support the hypothesis that high reversal rates in the conservative–old condition are related to liberal response biases.

Our finding that response reversals were more frequent following conservative than liberal conditions conflict with the results of Van Zandt and Maldonado-Molina (2004), who found similar reversal rates for liberal and conservative tests. Unlike the present experiment, their subjects tended to respond somewhat conservatively overall. Perhaps this was related to the use of foot pedals for the initial responses in which “no” responses were always assigned to the right foot. Subjects may have had a natural preference for their right foot because of their experience with driving cars, although Van Zandt has indicated that their pedals did not feel like car pedals (personal communication, December 9, 2004). Judging by d' for second responses, discrimination was higher in Van Zandt and Maldonado-Molina's Experiment 1 ($d' \approx 1.65$) than in the current experiment ($d' = 1.24$), so higher discrimination levels could explain their more conservative responses (Hirshman, 1995; Stretch & Wixted, 1998). Overall, our subjects were much more likely to reverse their responses (38% vs. ~18% for error reversals), so the d' difference between first and second responses was much larger in the current experiment ($d' = 0.78$ vs. 1.12) than for Van Zandt and Maldonado-Molina ($d' \approx 1.55$ vs. 1.65). Finally, it should be noted that Van Zandt and Maldonado-Molina's Figure 3 (p. 1153) appears to show a trend for more error reversals in conservative conditions, like the present results, although the effect was reported to be nonsignificant. More research will be needed to examine the many procedural differences that could have led to the different results. Virtually all parameters of the methods differed including list length, study duration, number of blocks, and method of inducing criterion shifts (Van Zandt & Maldonado-Molina varied the proportion of studied item in each test list rather than using payoffs).

General Discussion

The general purpose of the present experiments was to investigate criterion setting and response conflict during recognition

memory. Experiment 1 manipulated decision criteria by varying the rewards for accurate responses and recorded ERP activity when subjects used a liberal or conservative response bias. Experiment 2 was conducted to replicate the results of Experiment 1 and to investigate the amount of response conflict associated with the various types of memory judgments. Experiment 3 further investigated the hypothesis that old responses in the conservative condition were associated with a greater degree of conflict than all of the other responses by measuring response reversals when fast initial judgments were followed by slower judgments.

Conflict

Across Experiment 1 and Experiment 2, manipulating response criteria with asymmetric payoffs consistently influenced the 1,000- to 1,500-ms LPN such that the LPN was most negative following correct “old” responses during recognition tests with conservative criterion setting. Experiment 2 results indicate that the stimulus-locked LPN corresponds to the response-locked ERN, which was also largest in the conservative–old condition. The present ERP results are consistent with the argument that the LPN–ERN observed in recognition memory experiments reflects response conflict due to conflicting sources of information (Johansson & Mecklinger, 2003; R. Johnson et al., 2004, 2005). In the present case, conservative–old trials represent cases in which there were potential conflicts between the instructions to respond conservatively and the information retrieved from memory indicating that an item had been previously studied. A more thorough assessment of this interpretation would include analysis of error trials with the prediction that LPN–ERN enhancement would be observed only for hits and not misses because “new” responses associated with misses do not conflict with conservative incentives to respond “new.” Unfortunately, error rates were too low to support such analyses in the present experiments. Future experiments could examine this possibility by increasing error rates or test trials. However, our conflict interpretation does not rest on the ERP results alone. Experiment 3 provided behavioral evidence for response conflict by showing that subjects were most likely to reverse the responses to old items in the conservative condition. Thus, the present experiments provide clear support for the hypothesized relationship between the LPN–ERN and response conflict.

An apparent difficulty with our conflict interpretation concerns the fact that the LPN–ERN was enhanced in one conflicting situation (hits in conservative blocks) but not in another (correct rejections in liberal blocks). We hypothesized that the greater conflict in the conservative–old condition may have resulted because the inducements in the conservative condition moved subjects away from their natural tendency to use a more liberal criterion to make recognition judgments. Experiment 3 tested this hypothesis by measuring response bias when the rewards were equivalent for correct old and new judgments as well as allowing subjects to reverse their initial old–new judgment. Results from this experiment were consistent with the hypothesis that subjects in these experiments had a natural proclivity to respond liberally. Given the shallow encoding task of syllable counting, it is likely that the low subjective memorability of the lists induced subjects to naturally adopt a liberal criterion (Hirshman, 1995; Stretch & Wixted, 1998). Furthermore, when given the opportunity to re-

verse their first quick response, subjects were more likely to reverse in conservative than in liberal blocks. However, contrary to the hypothesized relationship between liberal response tendencies and more prevalent response reversals in conservative conditions, we failed to observe a significant correlation between response bias and response reversals. Although we hesitate to completely dismiss this hypothesis on the basis of null correlations across 20 subjects, alternative views should be considered.

It also is possible that conservative–old conditions were associated with greater conflict than liberal–new conditions because qualitatively different factors drive “old” versus “new” judgments. If an old item is highly familiar and/or clearly recollected, there exists a positive source of information that could directly conflict with the incentive to favor “no” responses in the conservative condition. In the absence of such positive evidence, correct rejection is often thought to rely on metacognitive strategies including assessment of the subjective memorability of items and the effectiveness of the encoding and retrieval conditions (Brown, Lewis, & Monk, 1977; Gallo, Weiss, & Schacter, 2004; Ghetti, 2003; Strack & Bless, 1994). Perhaps there is less opportunity for conflict with new test items because they are less likely to elicit the retrieval of information from memory that could potentially contradict the liberal test instructions.

Although we have primarily interpreted our LPN–ERN results as related to conflict, it should be noted that our results also may be compatible with an error signaling account of the ERN (e.g., Frank et al., 2005; Holroyd & Coles, 2002). ERNs are most often recorded under conditions in which errors are either subjectively easy to detect or are signaled by feedback (Holroyd & Coles, 2002). In the present experiments, errors were not subjectively obvious and feedback was not provided, so subjects’ subjective assessment may be more relevant than objective accuracy. When errors were increased by speeded responding, this error signal may have led to observable response reversals (Experiment 3). When accuracy was high and ERP analyses were limited to only accurate trials (Experiments 1 and 2), the error signal may reflect a subjective assessment that some responses may have been wrong, and thereby led to second-guessing of responses that were objectively correct.

Criterion Setting

There has been considerable recent interest in the factors controlling the setting of decision criteria in recognition memory tests. Evidence suggests that memory strength influences decision criteria such that conditions associated with higher memorability elicit more conservative criteria. Such strength-based criterion shifts are readily observed when strength is manipulated between different memory lists (Hirshman, 1995; Stretch & Wixted, 1998; Verde & Rotello, in press). Subjects often seem reluctant to shift criteria from item to item within a single list (Morrell, Gaitan, & Wixted, 2002; Stretch & Wixted, 1998; Verde & Rotello, in press), although such effects have recently been reported (Benjamin & Bawa, 2004; Dobbins & Kroll, 2005; Singer & Wixted, 2006).

The present ERN results are likely to reflect a postresponse consequence of criterion setting rather than the criterion setting process itself. The observation that subjects rarely shift criteria from item-to-item is consistent with a standard signal detection perspective in which subjects set their response criterion at the

beginning of a trial or block of trials prior to assessing an item’s memorability (or familiarity or retrieval strength). Response criteria placement may be determined by factors such as confidence or payoffs. From this perspective, it would be expected that our subjects would have set their criteria at the beginning of a test block on the basis of the specified payoffs. Thus, one might predict the consequence of different criterion settings to be observed relatively early in the ERP epoch. As reviewed in the introduction of this article, the most prominent memory-related ERP effects occur between about 300 and 800 ms following onset of a recognition test item and are thought to be related to core memory retrieval processes (reviewed by Curran, Tepe, & Piatt, 2006), but these standard ERP old–new effects were not reliably influenced by the present payoff manipulations. Our response bias effects were primarily limited to the 1,000- to 1,500-ms period that has been generally related to postretrieval control processes. Thus, the present results are more consistent with the perspective whereby criterion setting reflects the activity postretrieval metacognitive control processes that are important for the evaluation of memory (Benjamin & Bawa, 2004; Dobbins & Kroll, 2005; Gallo et al., 2004; Ghetti, 2003; Strack & Bless, 1994; Strack & Forster, 1995; Whittlesea, 1993; Whittlesea, Jacoby, & Girard, 1990). More precisely, our ERN results were observed following the subject’s response and so are more likely to reflect a later consequence of the chosen criterion rather than reflecting criterion setting per se. These postresponse consequences may be related either to the subject’s ongoing assessment of memory after their initial response or to the subject’s possible criterion adjustment on later trials. Each of these possibilities will be discussed in turn.

One possible postresponse assessment may be related to confidence. Van Zandt (2000) proposed a race model of recognition whereby evidence accumulates separately for “new” and “old” judgments. Confidence reflects the balance of evidence between these counters, such that judged confidence increases with the difference in evidence for “new” versus “old” judgments. According to the model, a highly confident “old” judgment occurs when the old counter has accumulated much more evidence than the new counter. A less confident “old” response occurs when the old counter is just slightly higher than the new counter. For the present purposes, the most important aspect of this perspective is that confidence does not reflect a preretrieval criterion setting process, but rather reflects a postretrieval assessment based on the information available at the time of the judgment. On the basis of results from a dual-response procedure similar to the present Experiment 3, the model was revised to allow for the continued accumulation of evidence after the initial response (Van Zandt & Maldonado-Molina, 2004). The present ERP results suggest that the LPN–ERN may provide evidence for the postresponse continuation of information processing under more natural conditions involving single responses given at typical speeds.

However, other ERP evidence has been consistent with confidence effects arising from preretrieval criterion setting. Azimian-Faridani and Wilding’s (2006) results showing confidence-based criterion effects on the 300- to 500-ms FN400 old–new effect are consistent preretrieval criterion perspective. Although our payoff-based criterion manipulations did not replicate the confidence-based manipulations of Azimian-Faridani and Wilding, other evidence from our lab has provided more consistent results by

measuring subjects' confidence ratings. When subjects rated the confidence of their responses, FN400 old–new differences (on correct trials) were larger for high than for low confidence responses (see Figure 14a in Curran, 2004). From the criterion setting perspective, high confidence responses correspond to extremely high (for old items) versus extremely low (for new items) response criteria, so the relative familiarity of hits and correct rejections would diverge more for high than low confidence responses.

In addition to considering how ERN–LPN activity in the present experiments may be related to control processes relevant to the assessment of the current item, it is possible that these control processes are relevant to the processing of future items. In domains outside of recognition memory, various mechanisms have been considered whereby the evaluation of earlier trials leads to changes in response criteria for later trials. These mechanisms might include monitoring stimulus changes (Brown & Steyvers, 2005), response monitoring (Treisman & Williams, 1984), and error-rate monitoring (Rabbitt, 1981). The latter two possibilities are particularly intriguing given that the ERN is response locked and has been associated with error monitoring (e.g., Frank et al., 2005; Holroyd & Coles, 2002).

Correlations between ERN amplitude and response bias changes in Experiment 2 were consistent with the possibility that the ERN may reflect the activity of an error-response monitoring mechanism related to criterion shifts. We calculated response bias (c) for each fourth of the sequentially ordered trials within each condition. Across subjects, the size of the c difference between the first and last fourth of trials correlated with ERN amplitude ($r = .36, p < .05$). Subjects with larger ERNs tended to become more liberal across trials, whereas those with smaller ERNs tended to become more conservative across trials. The correlation was not significant for liberal blocks ($r = -.04$). Response bias changes cannot be interpreted unambiguously as criterion shifts because they could also reflect movement in the underlying memory strength distributions (Wixted & Stretch, 2000), but these results are nonetheless highly suggestive of criterion shifts being related to ERN amplitude.

Conclusions

The results of these experiments suggest several conclusions. First, memory-related information processing does not stop when a judgment is made. This can be demonstrated behaviorally when subjects correct fast errors (present Experiment 3; Van Zandt & Maldonado-Molina, 2004) as well as electrophysiologically following accurate responses in the absence of speeded responding (present Experiments 1 and 2). Second, both ERN–LPN amplitudes and behavioral response reversals were greatest in the conservative–old conditions, as would be expected from the perspective that the LPN–ERN indexes response conflict associated with memory judgments (Johansson & Mecklinger, 2003). Third, payoff-induced response bias differences were associated with relatively late ERP effects (1,000- to 1,500-ms poststimulus LPN and 60-ms postresponse ERN). The late timing of these effects is consistent with payoffs influencing postresponse control process that may be related to either the subject's ongoing assessment of memory after their initial response or to criterion adjustment on later trials.

References

- Allan, K., Wilding, E. L., & Rugg, M. D. (1998). Electrophysiological evidence for dissociable processes contributing to recollection. *Acta Psychologica, 98*, 231–252.
- Azimian-Faridani, N., & Wilding, E. L. (2006). The influence of criterion shifts on electrophysiological correlates of recognition memory. *Journal of Cognitive Neuroscience, 18*, 1075–1086.
- Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., & Gratton, G. (2005). Strategic control and medial frontal negativity: Beyond errors and response conflict. *Psychophysiology, 42*, 33–42.
- Benjamin, A. S., & Bawa, S. (2004). Distractor plausibility and criterion placement in recognition. *Journal of Memory and Language, 51*, 159–172.
- Brown, J., Lewis, V. J., & Monk, A. F. (1977). Memorability, word frequency, and negative recognition. *Quarterly Journal of Experimental Psychology, 29*, 461–473.
- Brown, S., & Steyvers, M. (2005). The dynamics of experimentally induced criterion shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 587–599.
- Bunge, S. A., Burrows, B., & Wagner, A. D. (2004). Prefrontal and hippocampal contributions to visual associative recognition: Interactions between cognitive control and episodic retrieval. *Brain and Cognition, 56*, 141–152.
- Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition, 28*, 923–938.
- Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia, 42*, 1088–1106.
- Curran, T., & Cleary, A. M. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Cognitive Brain Research, 15*, 191–205.
- Curran, T., DeBuse, C., Worocho, B., & Hirshman, E. (2006). Combined pharmacological and electrophysiological dissociation of familiarity and recollection. *Journal of Neuroscience, 26*, 1979–1985.
- Curran, T., & Dien, J. (2003). Differentiating amodal familiarity from modality-specific memory processes: An ERP study. *Psychophysiology, 40*, 979–988.
- Curran, T., & Friedman, W. J. (2003). Differentiating location- and distance-based processes in memory for time: An ERP study. *Psychonomic Bulletin & Review, 10*, 711–717.
- Curran, T., & Friedman, W. J. (2004). ERP old/new effects at different retention intervals in recency discrimination tasks. *Cognitive Brain Research, 8*, 107–120.
- Curran, T., Schacter, D. L., Johnson, M. K., & Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. *Journal of Cognitive Neuroscience, 13*, 201–216.
- Curran, T., Tanaka, J. W., & Weiskopf, D. M. (2002). An electrophysiological comparison of visual categorization and recognition memory. *Cognitive, Affective, & Behavioral Neuroscience, 2*, 1–18.
- Curran, T., Tepe, K. L., & Piatt, C. (2006). ERP explorations of dual processes in recognition memory. In H. D. Zimmer, A. Mecklinger, & U. Lindenberger (Eds.), *Binding in human memory: A neurocognitive approach* (pp. 467–492). Oxford, United Kingdom: Oxford University Press.
- Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science, 5*, 303–305.
- Dien, J. (1998). Issues in the application of the average reference: Review, critiques, and recommendations. *Behavior Research Methods, Instruments and Computers, 30*, 34–43.
- Dobbins, I. G., & Kroll, N. E. A. (2005). Distinctiveness and the recognition mirror effect: Evidence for an item-based criterion placement heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1186–1198.

- Frank, M. J., Worocho, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*, 495–501.
- Friedman, D., Cycowicz, Y. M., & Bersick, M. (2005). The late negative episodic memory effect: The effect of recapitulating study details at test. *Cognitive Brain Research*, *23*, 185–198.
- Gallo, D. A., Weiss, J. A., & Schacter, D. L. (2004). Reducing false recognition with criterial recollection tests: Distinctiveness heuristic versus criterion shifts. *Journal of Memory and Language*, *51*, 473–493.
- Ghetti, S. (2003). Memory for nonoccurrences: The role of metacognition. *Journal of Memory and Language*, *48*, 722–739.
- Glanzer, M., Kim, K., Hilford, A., & Adams, J. K. (1999). Slope of the receiver operating characteristic in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 500–513.
- Harvey, L. O., Jr. (2005). Parameter estimation of signal detection models: RscorePlus (Version 5.5.2) [Computer software and manual]. Retrieved from <http://psych.colorado.edu/~lharvey/html/software.html>
- Healy, A. F., & Kubovy, M. (1978). The effects of payoffs and prior probabilities on indices of performance and cutoff location in recognition memory. *Memory & Cognition*, *6*, 544–553.
- Hirshman, E. (1995). Decision processes in recognition memory: Criterion shifts and the list-strength paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 302–313.
- Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709.
- Holroyd, C. B., Dien, J., & Coles, M. G. (1998). Error-related scalp potentials elicited by hand and foot movements: Evidence for an output-independent error-processing system in humans. *Neuroscience Letters*, *242*, 65–68.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.
- Johansson, M., & Mecklinger, A. (2003). The late posterior negativity in ERP studies of episodic memory: Action monitoring and retrieval of attribute conjunctions. *Biological Psychology*, *64*, 91–117.
- Johnson, M. K., Kounios, J., & Nolde, S. F. (1996). Electrophysiological brain activity and memory source monitoring. *NeuroReport*, *7*, 2929–2932.
- Johnson, R., Jr., Barnhardt, J., & Zhu, J. (2003). The deceptive response: Effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. *Biological Psychology*, *64*, 217–253.
- Johnson, R., Jr., Barnhardt, J., & Zhu, J. (2004). The contribution of executive processes to deceptive responding. *Neuropsychologia*, *42*, 878–901.
- Johnson, R., Jr., Barnhardt, J., & Zhu, J. (2005). Differential effects of practice on the executive processes used for truthful and deceptive responses: An event-related brain potential study. *Cognitive Brain Research*, *24*, 386–404.
- Junghöfer, M., Elbert, T., Tucker, D. M., & Braun, C. (1999). The polar average reference effect: A bias in estimating the head surface integral in EEG recording. *Clinical Neurophysiology*, *110*, 1149–1155.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Leynes, P. A., Bink, M. L., Marsh, R. L., Allen, J. D., & May, J. C. (2003). Test modality affects source monitoring and event-related potentials. *American Journal of Psychology*, *116*, 389–413.
- Luu, P., Tucker, D. M., Derryberry, D., Reed, M., & Poulsen, C. (2003). Electrophysiological responses to errors and feedback in the process of action regulation. *Psychological Science*, *14*, 47–53.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. (2nd ed.). Mahwah, NJ: Erlbaum.
- Malmberg, K. J., Holden, J. E., & Shiffrin, R. M. (2004). Modeling the effects of repetitions, similarity, and normative word frequency on old-new recognition and judgments of frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 319–331.
- Maratos, E. J., Allan, K., & Rugg, M. D. (2000). Recognition memory for emotionally negative and neutral words: An ERP study. *Neuropsychologia*, *38*, 1452–1465.
- Morrell, H. E., Gaitan, S., & Wixted, J. T. (2002). On the nature of the decision axis in signal-detection-based models of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1095–1110.
- Nessler, D., & Mecklinger, A. (2003). ERP correlates of true and false recognition after different retention delays: Stimulus- and response-related processes. *Psychophysiology*, *40*, 146–159.
- Nessler, D., Mecklinger, A., & Penney, T. B. (2001). Event related brain potentials and illusory memories: The effects of differential encoding. *Cognitive Brain Research*, *10*, 283–301.
- Nessler, D., Mecklinger, A., & Penney, T. B. (2005). Perceptual fluency, semantic familiarity and recognition-related familiarity: An electrophysiological exploration. *Cognitive Brain Research*, *22*, 265–288.
- Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary learning systems approach. *Psychological Review*, *110*, 611–646.
- Olichney, J. M., Van Petten, C., Paller, K. A., Salmon, D. P., Iragui, V. J., & Kutas, M. (2000). Word repetition in amnesia: Electrophysiological measures of impaired and spared memory. *Brain*, *123*, 1948–1963.
- Picton, T. W., Lins, O. G., & Scherg, M. (1995). The recording and analysis of event-related potentials. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 10, pp. 3–73). Amsterdam: Elsevier.
- Rabbitt, P. (1981). Sequential reactions. In D. H. Holding (Ed.), *Human skills* (pp. 153–175). London: Wiley.
- Ranganath, C., & Paller, K. A. (2000). Neural correlates of memory retrieval and evaluation. *Brain Research: Cognitive Brain Research*, *9*, 209–222.
- Ratcliff, R., Shue, C.-F., & Gronland, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, *99*, 518–535.
- Reder, L. M., Nhouyvanisong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 294–320.
- Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998, April 9). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*, 595–598.
- Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: A cognitive neuroscience perspective. *Trends in Cognitive Sciences*, *7*, 313–319.
- Scheffers, M. K., & Coles, M. G. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 141–151.
- Singer, M., & Wixted, J. T. (2006). Effect of delay on recognition decisions: Evidence for a criterion shift. *Memory & Cognition*, *34*, 125–137.
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Application to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*, 34–50.
- Srinivasan, R., Nunez, P. L., Silberstein, R. B., Tucker, D. M., & Cadusch, P. J. (1996). Spatial sampling and filtering of EEG with spline-Laplacians to estimate cortical potentials. *Brain Topography*, *8*, 355–366.
- Strack, F., & Bless, H. (1994). Memory for nonoccurrences: Metacognitive and presuppositional strategies. *Journal of Memory and Language*, *33*, 203–217.

- Strack, F., & Forster, J. (1995). Reporting recollective experiences: Direct access to memory systems? *Psychological Science*, *6*, 352–358.
- Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1379–1396.
- Treisman, M., & Williams, T. C. (1984). A theory of criterion setting with an application to sequential dependencies. *Psychological Review*, *91*, 68–111.
- Tucker, D. M. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography and Clinical Neurophysiology*, *87*, 154–163.
- Tucker, D. M., Luu, P., Frishkoff, G., Quiring, J., & Poulsen, C. (2003). Frontolimbic response to negative feedback in clinical depression. *Journal of Abnormal Psychology*, *112*, 667–678.
- van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & Behavior*, *77*, 477–482.
- Van Zandt, T. (2000). ROC curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 582–600.
- Van Zandt, T., & Maldonado-Molina, M. M. (2004). Response reversals in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1147–1166.
- Verde, M. F., & Rotello, C. M. (in press). Memory strength and the decision process in recognition memory. *Memory & Cognition*.
- Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1235–1253.
- Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language*, *29*, 716–732.
- Wilding, E. L. (1999). Separating retrieval strategies from retrieval success: An event-related potential study of source memory. *Neuropsychologia*, *37*, 441–454.
- Wilding, E. L., & Rugg, M. D. (1997a). An event-related potential study of recognition memory for words spoken aloud or heard. *Neuropsychologia*, *35*, 1185–1195.
- Wilding, E. L., & Rugg, M. D. (1997b). Event-related potential and the recognition memory exclusion task. *Neuropsychologia*, *35*, 119–128.
- Windmann, S., & Kruger, T. (1998). Subconscious detection of threat as reflected by an enhanced response bias. *Consciousness and Cognition*, *7*, 603–633.
- Windmann, S., & Kutas, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience*, *13*, 577–592.
- Windmann, S., Urbach, T. P., & Kutas, M. (2002). Cognitive and neural mechanisms of decision biases in recognition memory. *Cerebral Cortex*, *12*, 808–817.
- Wixted, J. T., & Stretch, V. (2000). The case against a criterion-shift account of false memory. *Psychological Review*, *107*, 368–376.
- Yeung, N., Cohen, J. D., & Botvinick, M. M. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931–959.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517.
- Yovel, G., & Paller, K. A. (2004). The neural basis of the butcher-on-the-bus phenomenon: When a face seems familiar but is not remembered. *NeuroImage*, *21*, 789–800.

Received January 12, 2006

Revision received June 30, 2006

Accepted July 25, 2006 ■