Event-related potentials during a short-term memory task: The role of stimulus-type

Jared Danker
Volen Center for Complex Systems, Brandeis University

Abstract
Dual-process theorists posit that two components of the event-related potential (ERP), the N400 and the P300, index the processes of familiarity and recollection, respectively. We recorded EEG while subjects performed a Sternberg (1966) memory task using verbal stimuli (letters, words, and objects) and non-verbal stimuli (spatial locations and sinusoidal gratings). We found that verbal and non-verbal stimuli elicit qualitatively different ERPs during encoding and retrieval. The N400 differentiates old and new verbal stimuli, but does not differentiate old and new non-verbal stimuli. However, the P300 differentiates old and new stimuli of both types. We conclude that familiarity requires a semantic label, but recollection does not.

According to dual-process theorists, dissociations observed in recognition memory tasks can be explained by two separable memory processes: recollection and familiarity (for review, see Yonelinas, 2002). Dual process theorists generally describe recollection as a process involving the recognition of specific details, including episodic information, while they describe familiarity as a feeling of knowing in the absence of source information. The classic illustration of this dissociation is the experience of being “familiar” with a person’s face without being able to “recollect” the person’s name (Mandler, 1991). Despite years of research, dual-process theorists still lack a full understanding of how familiarity and recollection operate, and in what ways these processes differ from each other. In the following study, we explore the differences in the event-related potentials (ERPs) during verbal and visual Sternberg (1966, 1975) memory tasks. We are interested in seeing how established effects from ERP studies of verbal memory generalized to visual memory, and most importantly, what this means for the current dual-process accounts of these effects.

The sizeable literature on the effect of stimulus repetition on the ERP has found that ERPs evoked by previously experienced stimuli are more positive-going than ERPs evoked by new stimuli (for review, see Johnson, 1995; Rugg, 1995). This effect, variably referred to as the “ERP repetition effect” or the “ERP old/new effect,” has been divided into early and late factors corresponding to the N400 and P300 components of the ERP. This division is justified by evidence demonstrating that these two components are sensitive to different experimental manipulations (Smith, 1993; Curran, 2000) and arise from distinct neural sources (Smith & Halgren, 1989; Guillem, N’Kaoua, Rougier, & Claverie, 1995). While initial attempts to interpret these two components within a dual-process framework
yielded a hodgepodge of relatively unsupported claims (Potter, Pickles, Roberts, & Rugg, 1992; Wilding, Doyle, & Rugg, 1995; Wilding & Rugg, 1996, 1997; Rugg, Schloerscheidt, & Mark, 1998), recent studies using novel experimental paradigms, including the remember/know procedure (Tulving, 1985), have led to a more lucid picture in which the N400 indexes familiarity and the P300 indexes a recollective process (Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Curran, 2000; Curran, Tanaka, & Weiskopf, 2002; Curran & Cleary, 2003; Curran, 2004). When interpreted within a dual-process framework, the remember/know procedure generates responses corresponding to both recollection (“remember”) and familiarity (“know”). Several studies have found that the P300 component of the ERP old/new effect is associated with “remember” responses, while the N400 component is associated with “know” responses (Smith, 1993; Curran, 2004; Duzel et al., 1997). Curran and colleagues (2000, 2003) have corroborated these findings in a series of experiments using a variation on the study-test procedure in which subjects are expected to differentiate between studied stimuli and new stimuli that are sometimes very similar to the studied stimuli. They found that stimuli similar to those on the study list led to fluctuations in the N400 consistent with a feeling of familiarity, while only items identical to those on the study list caused fluctuations in the P300. This is consistent with the P300 being an index of recollection, a process that involves the recognition of details.

An alternative to dual-process perspectives of the ERP old/new effect is the “context integration hypothesis” (Rugg & Doyle, 1994; Mecklinger, 1998). Rugg and Doyle (1994) argued that the N400 component represents the degree to which a word can be integrated with its context. According to the context integration hypothesis, contextual integration is facilitated when a word re-occurs one or more times in the same context. Furthermore, Rugg and Doyle (1994) claimed that context integration is dependent on certain stimulus properties, particularly the ability of the stimulus to be represented as a semantic construct that they refer to as a “unitary code.” Because of this restraint, the context integration hypothesis predicts that non-verbal stimuli should not elicit the N400 component of the ERP old/new effect.

Consistent with the predictions of the context integration hypothesis, a series of studies have shown that a robust ERP old/new effect is elicited by words (Rugg, 1985), non-words (Rugg, 1987; Rugg & Nagy, 1987), and meaningful pictures (Zhang, Begleiter, Porjesz, & Litke, 1997; Rugg & Doyle, 1994), but no effect is evoked by various non-verbal stimuli, including orthographically illegal nonwords (Rugg & Nagy, 1987), meaningless pictures (Rugg & Doyle, 1994; Thomas, 1992; Zhang et al., 1997; Crites, Delgado, Devine, & Lozano, 2000), novel objects (Rugg, 1995; Beisteiner et al., 1996), and spatial locations (Bosch, Mecklinger, & Friederici, 2001; Mecklinger & Muller, 1996; Mecklinger, 1998; Mecklinger & Bosch, 1999; Mecklinger & Meinshausen, 1998). While two studies have found ERP old/new effects using seemingly non-verbal stimuli (Curran et al., 2002; Penney, Mecklinger, Hilton, & Cooper, 2000), both of these studies involved substantial training that may have allowed the subjects to categorize and semantically encode the stimuli.

Electrophysiological studies of recognition memory find that repeated words elicit more positive-going ERPs when compared to new words. The effect encompasses both the N400 and P300 components of the ERP and is often absent when non-verbal stimuli are used, but most investigations of this effect using non-verbal stimuli have not looked at the N400 and P300 components of the ERP separately. It is important to analyze these
ERPS AND MEMORY FOR DIFFERENT STIMULUS-TYPES

components separately given evidence that they arise from different sources (Guillem et al., 1995) and represent distinct processes (Curran, 2000). In the following study, five different stimulus-types are used in a Sternberg (1966) short-term scanning memory paradigm in an attempt to better understand the effect of stimulus rehearsability on the ERP and the ERP old/new effect in particular. Stimulus-types include letters, words, objects, spatial dot locations (Jonides et al., 1993), and sinusoidal gratings (Kahana & Sekuler, 2002; Zhou, Kahana, & Sekuler, 2004). It is our hope that by using a range of stimuli varying in several stimulus properties, including verbal rehearsability, we will acquire a better understanding of the early and late components of the ERP old/new effect, and whether or not they depend on a semantic code. Furthermore, better understanding the role that the N400 and P300 components play in visual and verbal recognition memory may in turn shed some light on the processes familiarity and recollection that have been associated with these components.\(^1\)

Method\(^2\)

Subjects

Subjects were 12 paid volunteers, ages 19 to 29. Eight of the subjects were male, four were female. All subjects were right-handed. Most subjects did not have a previous or present history of medical illness; one male subject reported childhood epilepsy that had long ceased. At the time of testing, all subjects were well rested, relaxed, were not suffering from allergies or colds, and had normal or corrected-to-normal vision. All subjects gave informed consent to a protocol reviewed and approved by the Brandeis Committee for the Protection of Human Subjects. Subjects were given a base payment of $120 plus a bonus payment proportional to their performance. This technique encouraged subjects to stay alert and interested.

Study Paradigm

Subjects participated in a total of five EEG–Sternberg study sessions that were scheduled on different days within a two-week period. Each subject was tested with stimuli from each of the five different stimulus–types (letter, word, object, spatial, grating). Stimuli were shown on a computer monitor. Every trial of the experiment was self-paced; subjects depressed an advance key; which was followed by a 400–ms interval before the onset of a fixation cue (an asterisk). Each trial started with a fixation cue centered in the middle of the computer monitor for a duration of 1 second (±200 ms jitter), followed by a study set of three stimuli. The inter-stimulus interval (ISI) was approximately 1050 ms. Each stimulus was shown on the monitor for 700 ms and disappeared for 275 ms ±75 ms random jitter. After the third stimuli disappeared from the monitor, a short retention interval (500 ms ±75 ms jitter) was followed by a fourth item (probe) for 750 ms. Subjects were instructed to determine as quickly and accurately as possible after probe onset whether the probe item was part of the preceding study–set. They were instructed to respond to a target (a stimulus previously presented in that trial’s study-set) by pressing the right control key (dominant hand); they were to respond to a lure (a stimulus not presented in the trial’s study-set) by pressing the left control key (non-dominant hand). After responding, subjects were given a

---

\(^1\)For a more thorough review of ERP old/new and repetition effects, see the Appendix.

\(^2\)Method written by Grace Hwang-Grodzins and Aaron Geller.
ERPS AND MEMORY FOR DIFFERENT STIMULUS-TYPES

blink break and prompted to continue at their own pace by pressing a pre-selected down-arrow key. To prevent potential interference between a completed trial and the start of the next trial, a minimum of 1.5 seconds was preprogrammed to pass before the next trial could begin. Feedback on accuracy and response time was given at the end of each block. The temporal jitter in each ISI served to decorrelate the physiological responses from successive stimuli presentations.

Stimuli

Five pools of different stimulus-types were employed: single letters (letter), pictures of objects from the Snodgrass (Snodgrass & Vanderwart, 1980) picture database (object), one-syllable words corresponding to the object pool (word), single dot that appear at different positions on the monitor (spatial), and grey-scale sinusoidal patterns (grating). Each stimulus-type pool contained sixteen stimuli. A stimulus is considered verbal if it could be subvocally rehearsed with a semantic label. A stimulus is considered visual if it resists subvocal rehearsal. Using this definition, letter, word, and object stimuli are verbal, dot and grating stimuli are visual. A sample of each stimulus-type is shown in Figure 1.

The letter pool contained the following letters: b, c, d, f, g, h, j, k, l, m, n, p, q, r, t, and v. The object pool contained pictures of the following nouns: ball, bat, bed, bell, cake, car, chair, dog, ear, fly, fork, hat, heart, key, kite, and shoe. Each word in the word pool corresponded exactly with each object in the object pool. The spatial pool contained presentations of a single dot at sixteen different locations around an invisible circle of constant diameter (10 cm). This radius was intended to fall within the peripheral vision of the subject. Dot placements were random, did not overlap on the monitor, and did not follow times on a clock face and therefore resisted verbal coding. The grating pool contained two-dimensional textures, similar to those used in prior studies (Zhou et al., 2004; Kahana & Sekuler, 2002). The two-dimensional luminance profiles of each of the grating stimulus-type can be described by equation 1.

\[
L(x, y) = L_{avg}\left[1 + \frac{1}{2}A\sin(\pi fx) + \frac{1}{2}A\sin(\pi gy)\right]
\]

where \(L_{avg}\) represents mean luminance (L); \(f\) and \(g\) represent the spatial frequency of the vertical and horizontal components, respectively. where \(A\) is defined by

\[
A = \frac{L_{maximum} - L_{minimum}}{L_{maximum} + L_{minimum}}
\]

The gratings were generated by a custom program written in C++ using the Xlib and Xpm libraries. Parameters used to generate the 16 grating stimuli used in this experiment were \(A = 0.25, f = g = e^m\), where \(m = \{2.1, 2.6, 3.1, 3.6\}\) cycles/degree. The luminance of the monitor was linearized using calibration routines from Brainard & Pelli’s (1997) Psychtoolbox (Pelli, 1997).

Stimuli Constraints

Each block of 30 trials comprised 15 targets and 15 lures. The target trials, in turn, probed each study-position with equal probability. The sequence of target trials and lure
trials was randomized. The stimuli employed in each trial were randomized but subject to the following constraints:

1. When selecting study items for trial $t$, stimuli employed in trials $t - 1$ and $t - 2$ were disallowed.

2. A stimulus was allowed to be employed as a lure only once per block.

3. Lures were exempted from rule 1, unless they had been used as a target. For example, the following sequence of stimuli and trials would be legal, as the lure in the second trial was a non-probed study item in the preceding trial:

<table>
<thead>
<tr>
<th>Trial #</th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Probe</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>b</td>
<td>r</td>
<td>d</td>
<td>r</td>
</tr>
<tr>
<td>$t + 1$</td>
<td>n</td>
<td>k</td>
<td>j</td>
<td>d</td>
</tr>
</tbody>
</table>

The following sequence was disallowed, however:
4. To reduce the difficulty of the grating stimulus, immediate (non diagonal) neighbors on the frequency matrix were not permitted to occur in the same trial. For example, in the grating trial illustrated in Figure 2 the grating labeled A in Figure 2 has been selected, but gratings B through E were disallowed.

Subject Questionnaire

At the conclusion of the five–day study sessions, the subjects were interviewed and debriefed by one of the experimenters, typically the monitor. During the debriefing, a questionnaire was administered. This questionnaire was designed to assess the strategies that the subjects used to complete the study. Subjects were asked to report strategy–relevant information in three sections. In section 1, subjects were asked open-ended questions regarding the types of strategies they used to complete each of the five tasks (i.e., letters, words, objects, dot, grating). In addition, we asked about the development of subjects’ strategies within and across study sessions and about the relative effectiveness of those strategies. In section 2, subjects were asked to rate, on a 5–point scale, how often they used visual imagery and verbal names or labels to complete each of the five tasks (1 = never, 5 = always). Finally, in section 3, subjects were asked to rate the difficulty of each task relative to the other tasks on a continuum from 1 to 5 (1 = easiest, 5 = hardest).

Behavioral Data Analysis

Group means and variances were obtained for reaction time (RT) and accuracy according to stimulus–type (letter, word, object, dot, grating), sessions day (1-5), and probe–type (lure, probe position 1, probe position 2, probe position 3). Error bars were computed based on within–subject confidence intervals (Loftus & Masson, 1994) at the \( p = 0.05 \) level.

EEG Recording

All EEG recordings were made at Brandeis University, Volen Center for Complex Systems. During EEG recording, subjects sat comfortably in an ergonomic office chair 57 cm away from a flat–screen computer monitor. Individuals were asked to adjust their eye level such that their line-of-sight was at the same level as the center of the monitor. Subjects were instructed to remain silent during the task, to minimize all body and eye movements (particularly blinks), and to take as many breaks between trials as necessary to maintain their concentration during the task. During the entire study, an experimenter quietly monitored the session from the back of the testing room. Lighting in the room was kept constant within and between subjects.

Recordings were obtained from 60 tin electrodes located in standard electrode positions embedded in an elastic cap (ElectroCap). Electrode locations corresponded to AF3, AFz, AF4, F3, F4, F7, F8, F9, Fz, F10, FT9, FT7, FC5 ,FC3, FC1, FCz, FC2, FC4, FC6, FT8, FT10, T3, C3, Cz, C4, T4, TP9, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, TP10, P9, P7, P5, P3, P1, Pz, P2, P4, P6, P8, P10, PO9, PO7, PO3, POz, PO4, PO8, PO10, O1, Oz, O2, O9, and Iz.
Figure 2. This figure illustrates the additional distance constraint on stimuli selection for each trial of the grating task. These constraints reduce the difficulty of the grating task. Each column represents the number of horizontal frequency and each row represents the number of vertical frequency; both of which follow the set $e^{2.1}, e^{2.5}, e^{3.1}, e^{3.6}$. Immediate (non–diagonal) neighbors on this frequency matrix were not permitted to occur in the same trial. For example, in a grating trial in which the grating labeled A has been selected, gratings B through E were disallowed.
EEG signals were amplified 10,000 times (Sensorium EPA6), the amplifier analog filters were set at 0.03 Hz (high pass) and 50 Hz (low pass). The input impedance of the EPA6 amplifier is 10 G-ohm, rendering this EEG system very robust to high electrode impedance (Picton et al., 2000). Analog-to-digital signal conversion was implemented with a 12–bit data acquisition card (National Instrument PCI-6071E) with ± 5V dynamic range. The overall system resolution was therefore 0.24 µV/bit. Recording digitization rate was set at 256 Hz, which was well above the Nyquist sampling limit for our frequency region of interest (i.e., 2–50 Hz). Amplified signals were then referenced to minimize line noise, electronic artifacts, and experimental artifacts, and subsequently all signals were digitally notch-filtered between 59 and 61 Hz to minimize 60–Hz line noise. A single computer, operating in a real-time Linux Debian platform with shared memory enabled, was used to simultaneously record EEG signals and run the Sternberg experiment; this permitted every event to be tracked by a single system clock.

Electrode impedances were brought to < 50 K-ohm, inter-electrode impedances were within 20 K-ohm; skin impedances (ground and reference) were brought to < 10 K-ohm. Any electrodes that exhibited poor electrical characteristics were turned off at the head–box (Sensorium). All EEG signals were recorded referentially using the right mastoid (or right ear lobe). EEG signals were digitally rereferenced to the average EEG signal recorded from all electrically sound electrodes. Only signals recorded from low-impedance (< 50 K-ohm) and electrically sound electrodes were included in the re-referencing. On any given session, no more than 5% of all electrodes had poor electrical contact and or high impedances; this percentage is within the acceptable limit set forth by the Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria (Picton et al., 2000).

Electrooculogram (EOG)

Six tin disc-electrodes were used to monitor EOG activity. Vertical eye movements from the right eye were measured from one electrode positioned below the right eye, and from FP2, an electrode affixed just above the right eyebrow. Similarly, vertical eye movements from the left eye were measured from one electrode positioned below the left eye, and from FP1, an electrode affixed just above the left eyebrow. A pair of electrodes were placed at the outer canthi of each eye to measure horizontal eye movements. Each pair of EOG were recorded bipolarly. Raw signals from each pair of EOG were used for automatic blink detection following

\[ f(t) = 0.5 \times f(t-1) + 0.6 \times (D(t)) - s(t-1) \]  
\[ s(t) = 0.975 \times s(t-1) + 0.025 \times (D(t)) \]

where,
- \( f \) = fast rejection threshold.
- \( s \) = slow rejection, \( s(1) \) was set to the mean of the first 10 time points at the start of each event.
- \( t \) = time point;
- \( D \) = bipolarely recorded signals in \( \mu \) volts.
If any one pair of EOG exceeded a fast rejection threshold of $|100 \mu \text{Volts}|$, the event (e.g., first study item presentation) that corresponded to the exceedance was excluded from analysis. Fewer than 6\% of all trials were excluded.

**EEG Data Analysis**

Accuracy and RT were recorded for each trial. RTs shorter than 200 ms or greater than 1.3 seconds were excluded from analysis, this amounted to 2.1\% of all trials. Electrodes that exhibited high impedances (>50 K–ohm, <1\%) or electrical instability (e.g., electrode popping, DC drifts) were also excluded from analysis, as were those used to record electrooculogram.

**General**

Data were analyzed offline in Matlab (Mathworks, Natick, MA) using a combination of custom scripts and scripts from the Kahana Computational Lab EEG Toolbox (http://memory.psych.upenn.edu/programming/eeg_toolbox.php). All statistical analysis were done in Matlab.

**Event Related Potential (ERP) Analysis**

EEG signals were first digitally rereferenced based on the average signals collected from electronically stable EEG electrode, and then digitally low–pass filtered at 40 Hz. ERPs were then baseline–corrected using the mean voltage from the interval (100 ms) before the onset of the fixation stimulus.

To account for subject-to-subject and day-to-day signal fluctuation, Z–transformed ERPs were constructed from EEG signals for each subject, stimulus–type, session day following

$$
Z(t) = \frac{\frac{1}{N}\sum_{i=1}^{N} V_i(t) - \frac{1}{T} \sum_{t' \in T} V_i(t')} - \frac{1}{NB} \sum_{i=1}^{N} \sum_{t'' \in B} V_i(t'')}{\sigma \left( \frac{1}{N} \sum_{i=1}^{N} V_i(B) \right)}
$$

where,

\( t = \) time in time points for duration of interest (e.g., 1000 ms);
\( V = \) voltage ($\mu \text{V}$);
\( N = \) total number of trials in selected condition;
\( B = \) time corresponding to 400 ms before and to the the onset of the fixation stimulus at the start of trial;
\( \text{Boldface B} \) represents the number of items in vector B;
\( \sigma = \) standard deviation ($\mu \text{V}$);
\( T = \) time corresponding to 100 ms before and to the start of each trial;
\( \text{Boldface T} \) represents the number of items in vector T.

Finally, we take the mean Z–transformed ERP across days for each subject and adjust each subject’s Z–transformed ERP ($Z(t)$) to start at 0 voltage by applying the following

$$
Z'(t) = Z(t) - \frac{1}{T} \sum_{i \in T} Z(t)_i
$$

where,
Results

Subject Questionnaire

In order to assess the strategies used by subjects, responses were collected about the degree to which subjects used visual and verbal strategies to remember each of the stimulus-types. Table 1 displays the $M$ and $SD$ responses across subjects for each stimulus-type. A score of five indicates that subjects used a strategy all the time for that stimulus-type, while a score of one indicates that they never used the strategy. Subjects appeared to prefer verbal rehearsal strategies for the letter, word, and object stimuli, but they preferred visual strategies for spatial stimuli. Paired $t$ tests confirmed these observations ($t(11) = 2.66, 3.42, 3.08, -4.82, p < .05$). However, subjects report using both strategies equally for grating stimuli ($t(11) = 0, p > .05$). In addition, subjects were asked about the difficulty of each task, with responses ranging from 1 (very easy) to 5 (very hard). The results are displayed in the third row of Table 1. Subjects reported that the Spatial and Grating tasks were more difficult than the letter, word, and object tasks. Bonferroni-corrected paired $t$ tests confirmed that subjects found the spatial task to be more difficult than the letter, word, and object tasks ($t(11) = -7.72, -5.55, -3.48, p < .05$). Subjects also reported that the grating task was significantly more difficult than the letter, word, and object tasks ($t(11) = -15.28, -15.11, -12.85, -3.13, p < .05$). The remaining comparisons were not significant ($p > .05$).

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Letter</th>
<th>Word</th>
<th>Object</th>
<th>Spatial</th>
<th>Grating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal ($M \pm SD$)</td>
<td>4.42 ± 1.24</td>
<td>4.25 ± 1.06</td>
<td>4.42 ± 0.79</td>
<td>1.67 ± 1.07</td>
<td>3.58 ± 0.90</td>
</tr>
<tr>
<td>Visual ($M \pm SD$)</td>
<td>2.42 ± 1.73</td>
<td>2.17 ± 1.34</td>
<td>3.08 ± 1.31</td>
<td>4.50 ± 1.00</td>
<td>3.58 ± 1.31</td>
</tr>
<tr>
<td>Difficulty ($M \pm SD$)</td>
<td>1.16 ± 0.39</td>
<td>1.42 ± 0.51</td>
<td>1.83 ± 0.83</td>
<td>3.21 ± 0.99</td>
<td>4.33 ± 0.65</td>
</tr>
</tbody>
</table>

The Effects of Stimulus-Type, Probe-Type, and Session Day

The effect of probe-type and stimulus-type on behavior is shown in Figure 3A. The y-axis on the left plot represents accuracy, and the y-axis on the right plots represents reaction time in milliseconds. The x-axis on both plots is the probe-type, where $L$ represents lures and $T_1, T_2, T_3$ represent targets from that appeared as the first, second, and third item on the study list, respectively. The results for letter, word, and object stimuli are plotted in different shades of red, and those for spatial and grating stimuli are plotted in shades of blue. For accuracy, there was a trend whereby subjects performed extremely accurately on letter, word, and object trials, but performed less accurately on spatial and grating trials. There
also appeared to be a recency effect on accuracy, with \( T_3 \) being answered more accurately than \( T_1 \), \( T_2 \), or \( L \), especially for grating trials. In order to confirm these observations, a 5(stimulus-type)×4(probe-type) repeated-measures ANOVA was performed with accuracy as the dependent variable. There was a significant main effect of stimulus-type on accuracy (\( F(4,44) = 99.65, p < .01 \)). In order to explore which of the five stimulus-types were different from the others, Bonferroni-corrected pairwise comparisons were performed for each of ten possible stimulus-type pairs. Spatial and grating trials were answered significantly less accurately than trials from the other three stimulus-types, and spatial trials were also answered significantly more accurately than grating trials (\( p < .01 \)). As predicted, letter, word, and object trials did not differ significantly from each other in accuracy (\( p > .05 \)). There was also a significant effect of probe-type on response accuracy (\( F(3,33) = 8.90, p < .01 \)). To explore the effect of probe-type on response accuracy, Bonferroni-corrected pairwise comparisons were performed comparing each of the four probe-types. In agreement with the observation of a recency effect, responses to \( T_3 \) were significantly more accurate than responses to \( L \), \( T_1 \), and \( T_2 \) (\( p < .01 \)). \( L \), \( T_1 \), and \( T_2 \) did not differ significantly from each other on accuracy (\( p > .05 \)). Furthermore, stimulus-type significantly interacted with probe-type (\( F(12,132) = 4.99, p < .01 \)), with grating stimuli showing stronger recency effects that the other stimulus-types.

As can be seen in the right panel of Figure 3A, there was a clear recency effect for reaction time for all stimulus-types, with \( T_3 \) being answered faster than \( L \), \( T_1 \), and \( T_2 \). In addition, reaction time appeared to be the same for letter, word, object, and spatial trials, but substantially longer for grating trials. In order to confirm these observations, a second 5(stimulus-type)×4(probe-type) repeated-measures ANOVA was performed with reaction
time as the dependent variable. There was also a significant main effect of stimulus-type on reaction time \((F(4,44) = 67.40, p < .01)\). To test which stimulus-types were significantly different from each other, Bonferroni-corrected pairwise comparisons were performed on each of the possible stimulus-type pairs. Grating trials were answered significantly slower than trials of the other four stimulus-types \((p < .01)\). In agreement with our observations, reaction times on trials from the letter, word, object, and spatial stimulus-types were not significantly different from each other \((p > .05)\). There was a significant main effect of probe-type on reaction time \((F(3,33) = 30.84, p < .01)\). Bonferroni-corrected pairwise comparisons were performed comparing the reaction time of the four probe-types to each other. Confirming the presence of a recency effect, responses to \(T_3\) were answered significantly faster than responses to \(L, T_1,\) and \(T_2\) \((ps < .05)\). \(L, T_1,\) and \(T_2\) did not differ from each other significantly in reaction time \((ps > .05)\). In addition, there was a significant interaction between stimulus-type and probe-type for reaction time \((F(12,132) = 3.80, p < .01)\).

In order to ensure that the effects obtained during ERP analysis are due to the intended manipulation (i.e., stimulus-type) and not due to differences in performance, an attempt was made to equalize the mean accuracy of each stimulus-type. This was done by selecting and analyzing only trials with certain across-subject accuracies. Selected trials had at least 5 repetitions across the 12 subjects. In addition, trials drawn from verbal stimulus-types (letter, word, and object) were included only if they did not have a mean accuracy of 1.0. Trials drawn from the spatial stimulus-type were included only if they did not have 1.0 mean accuracy or mean accuracy less than 0.7. Trials drawn from the grating stimulus-type were included only if the mean accuracy exceeded or was equal to 0.4. The behavioral data using only these selected trials are plotted in Figure 3B. Accuracy appears to be well equated across stimulus-types. To ensure that accuracy was equated across modalities, \(5(\text{stimulus-type}) \times 4(\text{probe-type})\) repeated-measures ANOVAs were performed with accuracy and reaction time as the dependent variables using only these trials. For accuracy, there was neither a significant main effect of day \((F(4,40) = 1.32, p > .05)\), nor a significant interaction effect between day and stimulus-type \((F(16,160) = 1.21, p > .05)\). However, for reaction time, there was a significant main effect of day \((F(4,40) = 11, p < .01)\), and day and stimulus-type significantly interacted \((F(16,160) = 2.72, p < .01)\). The effect of day was a significant linear trend, with subjects answering faster over days \((p < .01)\). An accuracy-corrected plot similar to that displayed in Figure 3B is shown in Figure 4B. Once again, it appears that trials from the different stimulus-types were equated on accuracy.

In order to explore the effect of learning on behavior, the effect of session day and stimulus-type are plotted together in Figure 4A. The y-axis on the left plot represents accuracy and the y-axis on the right plot represents reaction time. Session day is plotted on the x-axis. Again, the accuracy and reaction time for letter, word, and object stimuli are plotted in different shades of red, and those for spatial and grating stimuli are plotted in shades of blue. \(5(\text{stimulus-type}) \times 4(\text{probe-type})\) repeated-measures ANOVAs were performed on the accuracy (left) and reaction time (right) measures. For accuracy, there was neither a significant main effect of day \((F(4,40) = 1.32, p > .05)\) nor a significant interaction effect between day and stimulus-type \((F(16,160) = 1.21, p > .05)\). However, for reaction time, there was a significant main effect of day \((F(4,40) = 11, p < .01)\), and day and stimulus-type significantly interacted \((F(16,160) = 2.72, p < .01)\). The effect of day was a significant linear trend, with subjects answering faster over days \((p < .01)\).
The Effect of Stimulus-Type and Learning on Behavior

Figure 4. The effect of stimulus-type and learning on behavior. A. The mean accuracy (left) and reaction time for correct trials (right) are plotted for each of the five testing sessions (1-5) and each of the five stimulus-types (letter, word, object, spatial, and grating). B. Same as the plot in A, only the accuracy across stimulus-types is equalized. Error bars represent confidence intervals for within-subject designs (Loftus & Masson, 1994).

The Effect of Study-Probe Distance

Models of recognition memory have shown that false alarms are more likely when the physical distance between a lure probe (L) and the items on the study list is small (Kahana & Sekuler, 2002). In order to explore the effect of study-probe distance, the distance in physical space was calculated between each study item and L. In the case of grating stimuli, this distance was the euclidean distance in spatial frequency in the x- and y-dimensions. In the case of spatial stimuli, this distance was the angle between the two dots along the 10cm circle of presentation. Each L trial was then ranked according to the distance between the probe and the study item to which it was closest. Trials were then binned into quartiles, and the most “near” and “far” quartiles were compared on accuracy and reaction time measures using paired t tests. Reaction time was only analyzed for correctly answered trials. These results are shown in Figure 5A for grating stimuli and Figure 5B for spatial stimuli.

For both grating and spatial stimuli, a trend existed whereby far L trials were answered more quickly and accurately than near L trials. To test whether these trends were significant, paired t tests were performed. For grating trials, there was no significant difference in response accuracy between near and far L trials (t(11) = -0.89, p > .05). However, there was a significant difference in reaction time between near and far grating trials, with far L trials being rejected more quickly than near L trials (t(11) = 2.85, p < .05). There was no significant difference in either accuracy (t(11) = -0.01, p > .05) or reaction time (t(11) = 0.45 p > .05) for spatial trials.

Event-Related Potentials (ERPs)
ERPs AND MEMORY FOR DIFFERENT STIMULUS-TYPES

Figure 5. The effect of lure-study distance. A. For grating trials, the mean accuracy and reaction time averaged across 12 subjects are presented separately for the quartile of trials in which the lure was farthest from one of the study items (“Far”) and the quartile of trials in which the lure was closest to one of the study items (“Near”). B. For spatial trials, the mean accuracy and reaction time averaged across 12 subjects are presented separately for far and near lures. Error bars represent confidence intervals for within-subject designs (Loftus & Masson, 1994), *p < .05.

ERP during the Study Interval

To investigate the effect of stimulus-type on stimulus encoding, comparisons were made between the ERPs elicited by different stimulus-types during the study interval, because during this period subjects are encoding the study items for later retrieval. To reduce the number of comparisons, the ERPs elicited by each of the three study items ($S_1$, $S_2$, $S_3$) were averaged, resulting in one study ($S$) ERP for each subject and stimulus-type. In addition, ERPs elicited by letter, word, and object stimuli were averaged to best contrast verbal and non-verbal memory, as subjects reported preferring verbal rehearsal strategies for all three of these stimulus-types. The resulting verbal study ($S_{verbal}$) ERPs were compared to grating study ($S_{grating}$) ERPs and spatial study ($S_{spatial}$) ERPs for three mean amplitude bins: Early (0-300 ms), N400 (300-500 ms), and P300 (500-700 ms). These bins were chosen for complete coverage of the study period and to roughly correspond to the time period during which the N400 and P300 components of the ERP occur. At each electrode, two paired t tests were performed to compare (1) $S_{verbal}$ ERPs to $S_{grating}$ ERPs and (2) $S_{verbal}$ ERPs to $S_{spatial}$ ERPs. In order to quantify the type I error rate independently of assumptions based on the normality of the data, a nonparametric resampling procedure for two paired samples with 1,000 repetitions was executed for each t test performed. Each repetition consisted of randomly shuffling the data across electrodes and the variable of interest (in this case, stimulus-type) and performing a t test. The t statistic obtained from the unshuffled data was then compared to the t distribution obtained from the 1000 shuf-
Figure 6. The effect of stimulus-type on the ERP during encoding: Verbal versus Grating. A. Topographic plots displaying significant differences between \( S_{\text{verbal}} \) and \( S_{\text{grating}} \) ERPs during three time periods: Early (0-300 ms), N400 (300-500 ms), and P300 (500-700 ms). Shades of blue indicate that \( S_{\text{grating}} \) ERPs were significantly more positive relative to \( S_{\text{verbal}} \) ERPs. Red indicates that \( S_{\text{verbal}} \) ERPs were significantly more positive than \( S_{\text{grating}} \) ERPs. B. Topographic plots corresponding to those in A, only the accuracy across stimulus-types is equalized. C. The top plots are example waveforms drawn from electrodes Cz (left) and PO9 (right) for \( S_{\text{verbal}} \) (red) and \( S_{\text{grating}} \) (blue). The bottom plots are the corresponding difference waveforms with positive values meaning that the \( S_{\text{verbal}} \) ERP was more positive, and negative meaning that the \( S_{\text{grating}} \) ERP was more positive. Error bars represent the 95% confidence interval.
fled repetitions, and a $p$ value was obtained. This method also corrects for the number of electrodes tested because each $p$ value is calculated based on an empirical distribution from all electrodes.

The resulting $p$ values for the comparison between $S_{\text{verbal}}$ and $S_{\text{grating}}$ ERPs are plotted topographically in Figure 6A. White indicates no significance ($p > .05$), while shades of blue indicate that $S_{\text{grating}}$ ERPs were more positive than $S_{\text{verbal}}$ ERPs at this time period and electrode location ($p < .05$). Shades of red indicate that $S_{\text{verbal}}$ ERPs were more positive than $S_{\text{grating}}$ ERPs ($p < .05$). Darker shades of either color indicate a smaller $p$ value.

Before 300 ms, $S_{\text{grating}}$ ERPs were significantly more positive compared to $S_{\text{verbal}}$ ERPs in right frontocentral regions of the scalp, while $S_{\text{verbal}}$ ERPs were more positive in occipitoparietal regions. Between 300 and 500 ms, $S_{\text{grating}}$ ERPs were more positive bilaterally at frontocentral locations and $S_{\text{verbal}}$ ERPs were again more positive at occipitoparietal locations. During the P300 time period, both effects became lateralized, with $S_{\text{grating}}$ ERPs being more positive in right frontal areas and $S_{\text{verbal}}$ ERPs being more positive in left temporoparietal areas. So although the effects are maintained over time, they change in topography, initially showing small effects before 300 ms which become more robust between 300 and 500 ms and lateralize between 500 and 700 ms after stimulus presentation. Figure 6B demonstrates that the topographies of these effects as seen in Figure 6A are largely maintained after accuracy adjustments are made (see Behavioral section for details). Figure 6C displays two example waveforms, one drawn from a central electrode (Cz) where $S_{\text{grating}}$ ERPs were more positive, and one drawn from a occipitoparietal electrode (PO9), where $S_{\text{verbal}}$ ERPs were more positive. Plotted in the bottom row of panel C are the difference between the two waveforms. In both cases the differences appear to be robust, perpetuating for several hundred milliseconds.

The corresponding plots for the comparison between $S_{\text{verbal}}$ and $S_{\text{spatial}}$ ERPs are displayed in Figure 7. In Figure 7A, white again reflects areas of no significance, while red indicates that $S_{\text{verbal}}$ ERPs were more positive than $S_{\text{spatial}}$ ERPs and blue indicates that $S_{\text{spatial}}$ ERPs were more positive. The topographies are highly similar to those displayed in Figure 6A, with a positivity in the $S_{\text{spatial}}$ ERP relative to verbal study ERP in frontocentral regions becoming more distributed and robust over time and eventually lateralizing to the right. Furthermore, the $S_{\text{verbal}}$ ERP was more positive than $S_{\text{spatial}}$ ERPs in occipitoparietal regions, but this effect fades to a large degree after 500 ms. Again, the topographies are largely maintained when the data is reanalyzed with accuracy being equated across the different stimulus-types (see Figure 7B). Example waveforms from central (Cz) and occipitoparietal (PO9) electrodes are displayed in Figure 7C.

**ERP during the Probe Interval**

*Old/New Effects.* To investigate the effect of memory on the probe ERP, items that were previously presented in the study list ($T$ = targets/old items) and items that were not presented in the study list ($L$ = lures/new items) were compared during the two time components that compose the ERP old/new effect: the early N400 (350-400 ms) and the late P300 (500-600 ms). These time windows were chosen in preference to wider windows because these two characteristic effects overlap in both time and space, but have distinct topographies. Therefore, conservative time windows were chosen for each component to best isolate its individual topography. Letter, word, and object probe ERPs were averaged to
Figure 7. The effect of stimulus-type on the ERP during encoding: Verbal versus Spatial. A. Topographic plots displaying significant differences between $S_{\text{verbal}}$ and $S_{\text{spatial}}$ ERPs during three time periods: Early (0-300 ms), N400 (300-500 ms), and P300 (500-700 ms). Shades of blue indicate that $S_{\text{spatial}}$ ERPs were significantly more positive relative to $S_{\text{verbal}}$ ERPs. Red indicates that $S_{\text{verbal}}$ ERPs were significantly more positive than $S_{\text{spatial}}$ ERPs. B. Topographic plots corresponding to those in A, only the accuracy across stimulus-types is equalized. C. The top plots are example waveforms drawn from electrodes Cz (left) and PO9 (right) for $S_{\text{verbal}}$ (red) and $S_{\text{spatial}}$ (blue). The bottom plots are the corresponding difference waveforms with positive values meaning that the $S_{\text{verbal}}$ ERP was more positive, and negative meaning that the $S_{\text{spatial}}$ ERP was more positive. Error bars represent the 95% confidence interval.
obtain a “verbal probe ERP.” This was justified both by similarities in the behavioral data as well as subject reports. At both time windows, a 3(stimulus-type) × 2(probe-type) repeated-measured ANOVA was performed at each electrode in order to best visualize the topography of the ERP old/new effect as well as the effect of stimulus-type and the interaction between these two effects. Again, in order to estimate type I error, a nonparametric resampling procedure with 1000 repetitions was executed for every ANOVA performed, with random shuffling across electrodes, stimulus-type, and probe-type.

The p values from the resampled ANOVAs are presented topographically in Figure 8. Panel A presents the main effects of stimulus-type and probe-type during the N400 component of the ERP, as well as the interaction between the two. In agreement with differences observed during encoding, there was a significant main effect of stimulus-type at frontocentral and occipitoparietal electrodes (p < .05). There was no main effect of stimulus-type with a frontocentral topography suggestive of the N400 old/new effect (p > .05). However, there was a significant interaction between stimulus-type and probe-type at frontocentral locations (p < .05), suggesting that T and L ERPs were different during this period for some stimulus-types, but not others. Panel B demonstrates that the topographies were not largely affected by accuracy-correction; however, the main effect of stimulus-type and the interaction effect appear somewhat attenuated by the correction.

The p values from the resampled ANOVAs during the P300 component of the ERP are presented in Figure 8C. The three topographic plots represent the significance of the
Figure 9. The ERP old/new effect for verbal, spatial, and grating stimulus-types: N400 (350-400 ms). A. Topographic plots displaying significant differences between the N400 components of T ERPs (old) and L ERPs (new) for verbal, spatial, and grating stimuli. Shades of red indicate that T ERPs were significantly more positive relative to L ERPs. Blue indicates that L ERPs were significantly more positive than T ERPs. B. Topographic plots corresponding to those in A, only the accuracy across stimulus-types is equalized. C. The top plots are example waveforms drawn from electrode Fz for T (red) and L (blue) probes. The bottom plots are the corresponding difference waveforms with positive values meaning that the T ERP was more positive, and negative meaning that the L ERP was more positive. Dashed vertical lines represent the time period analyzed in A and B. Error bars represent the 95% confidence interval.

main effect of stimulus-type, the main effect of probe-type, and the interaction between stimulus-type and probe-type, respectively. There was a significant main effect of stimulus-type at frontal and parietal electrodes (p < .05), and a significant effect of probe-type at right centroparietal and left temporal electrodes. A significant interaction between stimulus-type and probe-type was overlapped with the right centroparietal effect of probe-type, suggesting that this effect was different for different stimulus types. Panel D demonstrates that the topographies were maintained but the significance was attenuated by accuracy correction.

In order to visualize the topography of the ERP old/new effect separately for each stimulus-type, a paired t test comparing T and L ERPs was performed at both time windows for each stimulus-type (verbal, spatial, and grating) and electrode. A nonparametric resampling procedure with 1000 repetitions was executed for each t test performed, with random shuffling across electrodes and probe-type. The p values from the resampled t tests performed on the N400 time window are displayed topographically in Figure 9A for verbal, spatial, and grating ERPs. White indicates areas of no significance (p > .05). Red indicates areas where T ERPs were significantly more positive than L ERPs (p < .05), while blue represents locations where L ERPs were significantly more positive (p < .05).
ERPS AND MEMORY FOR DIFFERENT STIMULUS-TYPES

AT

B

C

Verbal Spatial Grating Verbal Spatial Grating Verbal ERP, C4 Spatial ERP, C4 Grating ERP, C4

Figure 10. The ERP old/new effect for verbal, spatial, and grating stimulus-types: P300 (500-600 ms). A. Topographic plots displaying significant differences between the P300 components of T ERPs (old) and L ERPs (new) for verbal, spatial, and grating stimuli. Shades of red indicate that T ERPs were significantly more positive relative to L ERPs. Blue indicates that L ERPs were significantly more positive than T ERPs. B. Topographic plots corresponding to those in A, only the accuracy across stimulus-types is equalized. C. The top plots are example waveforms drawn from electrode C4 for T (red) and L (blue) stimuli. The bottom plots are the corresponding difference waveforms with positive values meaning that the T ERP was more positive, and negative meaning that the L ERP was more positive. Dashed vertical lines represent the time period analyzed in A and B. Error bars represent the 95% confidence interval.

The first plot in Figure 9A shows the significance topography of the old/new effect for verbal stimuli during the N400 time window. The topography is representative of what could be called the canonical frontal N400 (or “FN400”) effect, with T ERPs being more positive at frontal recording sites and L ERPs being more positive in more posterior regions (Curran, 2000). The corresponding old/new significance plots for spatial and grating stimuli do not show any areas where T ERPs are more positive than L ERPs, and have only scattered significance in the opposite direction. Figure 9B shows that after accuracy-correction, while there is still some frontal positivity for T ERPs and posterior positivity for L ERPs, the significance of the N400 old/new effect is greatly attenuated.

Figure 9C shows representative T and L ERPs from electrode Fz for verbal, spatial, and grating stimuli. For each stimulus-type, L ERPs were subtracted from T ERPs. The resulting difference waveforms are plotted in the bottom row of panel C. A significant difference between T and L ERPs is apparent during the N400 analysis window (350-400ms, marked by dashed vertical lines) for verbal stimuli, but not for either spatial or grating stimuli.

The topographic significance plots for the old/new effect during the P300 time window
are displayed in Figure 10A for all three stimulus-types. As with Figure 9A, white indicates areas of no significance \((p > .05)\), while red and blue indicate significantly more positive ERPs for \(T\) and \(L\) ERPs, respectively \((p < .05)\). The first plot represents the difference between \(T\) and \(L\) ERPs during the P300 time window for verbal items. \(T\) ERPs are more positive than \(L\) ERPs at centroparietal regions for verbal items, and this effect is lateralized to the right. For both spatial and grating stimuli, the effect is topographically more constrained and more lateralized to the right, showing significance at only a few electrodes on the right side of the head. Difficulty-correction appears to attenuate, but not abolish, the P300 old/new effect for verbal and grating stimuli (see Figure 10B).

Example waveforms from electrode C4 (the electrode immediately to the right of the primary central electrode, Cz) are plotted in Figure 10C for verbal, spatial, and grating ERPs. The difference waveforms resulting from subtracting \(L\) ERPs from \(T\) ERPs are plotted in the bottom row of panel C. The P300 old/new effect is clearly strongest for verbal ERPs, but it is significant in ERPs of all three stimulus-types.

The Effect of Serial Position. In order to better understand the dynamics of the old/new effect, we investigated the effect of serial position on the N400 (350-400 ms) and P300 (500-600 ms) components of the ERP. The amplitude of the N400 component of the ERP at electrode Fz for each of the four probe-types \((L, T_1, T_2,\) and \(T_3\)) and the three stimulus-types (verbal, spatial and grating) is shown in Figure 11A. The corresponding plot for the P300 component at electrode C4 is shown in Figure 11B. During the N400 time window, all three stimulus-types show a trend whereby as probe position increases so does N400 amplitude. The corresponding waveforms for the four probe-types are presented separately for each stimulus-type in Figure 11C. For verbal stimuli, it is apparent that during the N400 time bin (marked by dashed vertical lines) N400 amplitude increases from \(L\) to \(T_1\) to \(T_2\) to \(T_3\). N400 amplitude also increases from \(T_1\) to \(T_2\) to \(T_3\) for spatial and grating stimuli, but as expected from our old/new analysis, \(L\) ERPs do not have the smallest amplitude.

To test whether the trend in serial position was significant, a \(3\times 3\) repeated-measures ANOVA was performed at electrode Fz. There was a significant main effect of probe position on the amplitude of the N400 \((F(2,22) = 15.31, p < .01)\). However, there was no main effect of stimulus-type on N400 amplitude \((F(2,22) = 1.50, p > .05)\), and stimulus-type did not significantly interact with probe position \((F(4,44) = 2.09, p > .05)\).

It can be seen in Figure 11B that the P300 component of the ERP did not increase in amplitude at electrode C4 as serial position increased. A \(3\times 3\) repeated ANOVA was performed to test this observation. There was no significant main effect of serial position on the amplitude of the P300 \((F(2,22) = 0.09, p > .05)\). There was, however, a significant main effect of stimulus-type on P300 amplitude \((F(2,22) = 8.90, p < .01)\). Serial position did not interact significantly with stimulus-type \((F(4,44) = 0.20, p > .05)\). No further analyses were performed on P300 because the most significant electrode for the old/new effect did not show serial position effects.

To investigate the topography of the recency effect during the N400 component of the ERP, paired comparisons were performed comparing \(T_3\) ERPs to the average of the \(T_1\) and \(T_2\) ERPs \((T_{1+2})\). A paired \(t\) test comparing these two ERPs was performed for visual, spatial, and grating stimuli at every electrode. A nonparametric resampling procedure
ERPs and Memory for Different Stimulus Types

Figure 11. The effect of serial position on the ERP. A. The mean amplitude of the ERP during the N400 (350-400 ms) component for each of the four probe-types and for verbal, spatial, and grating stimuli at electrode Fz. Error bars represent confidence intervals for within-subject designs (Loftus & Masson, 1994). B. The mean amplitude of the ERP during the P300 (500-600 ms) component for each of the four probe-types and for verbal, spatial, and grating stimuli at electrode C4. C. Example waveforms drawn from electrode Fz for L, T1, T2, and T3 are plotted for verbal, spatial, and grating stimuli. Dashed vertical lines represent the time period analyzed in A. D. Topographic plots displaying significant differences between the N400 component of T3 ERPs and T1+2 ERPs for verbal, spatial, and grating stimuli. Shades of red indicate that T3 ERPs were significantly more positive relative to T1+2 ERPs. Blue indicates that T1+2 ERPs were significantly more positive than T3 ERPs. E. Topographic plots displaying significant differences between T1+2 ERPs and L ERPs for verbal, spatial, and grating stimuli. Shades of red indicate that T1+2 ERPs were significantly more positive relative to L ERPs. Blue indicates that L ERPs were significantly more positive than T1+2 ERPs.
with 1000 repetitions was executed for each t test performed, with random shuffling across electrodes and serial position (T₃ versus T₁+₂). The resulting significance topographies for verbal, spatial, and grating stimuli are presented in Figure 11D. For verbal stimuli, the T₃ ERPs were significantly more positive at left frontal electrodes relative to the T₁+₂ ERPs. T₁+₂ ERPs were more positive at occipitoparietal electrodes bilaterally. For spatial stimuli, T₃ ERPs were more positive than T₁+₂ ERPs at central electrode locations, and T₁+₂ ERPs are more positive at left occipitoparietal locations. For grating stimuli, the T₃ ERPs were more positive at left frontal electrodes, while the T₁+₂ ERPs were more positive at left occipitoparietal electrodes. No accuracy-adjusted topographic plots are shown because the accuracy-adjustment resulted in the rejection of most T₃ trials for verbal stimuli and most T₁ trials for grating stimuli.

In order to ensure that the ERP old/new effect for verbal items was not caused entirely by the increased positivity of T₃ ERPs, T₁+₂ ERPs were compared to L ERPs for verbal, spatial, and grating stimuli. Paired t tests comparing T₁+₂ ERPs to L ERPs were performed for each stimulus-type at each electrode. As was done for the other comparisons, a nonparametric resampling procedure with 1000 repetitions was executed for each t test performed, with random shuffling across electrodes and probe-type (old versus new). The resulting significance topographies are presented in Figure 11E. As expected, the canonical N400 old/new effect is not apparent for either spatial or grating stimuli. However, a lateralized N400 effect is apparent for verbal stimuli, with T₁+₂ ERPs being more positive at right frontal locations and L ERPs being more positive at left occipitoparietal locations.

Discussion

The Effect of Stimulus-Type

ERPs were observed to differ between different stimulus-types during the presentation of both study items (stimulus encoding) and probe items (stimulus recognition). The effect was strongest after 300 ms, with non-verbal stimuli (spatial and grating) eliciting more positive ERPs than verbal stimuli (letter, word, object) initially in bilateral frontocentral regions and later lateralizing to electrodes located on the right side of the scalp. Verbal ERPs were more positive than non-verbal ERPs initially at bilateral occipitoparietal electrode locations, but this effect lateralized to left electrodes after 500 ms. The observed anterior/posterior dissociation between verbal and non-verbal memory is consistent with studies by Mecklinger and colleagues (Bosch et al., 2001; Mecklinger & Pfeifer, 1996) that showed that the encoding and retention of object working memory involved an increased negative slow wave in frontal locations relative to spatial working memory, and that the encoding and retention of spatial working memory involved an increased negative wave in parietal locations relative to object working memory. Mecklinger and colleagues (Bosch et al., 2001; Mecklinger & Pfeifer, 1996) interpret the frontal negative wave as being involved in object memory, and the parietal negative wave as being involved with a visuospatial storage mechanism (Awh, Jonides, & Reuter-Lorenz, 1998). However, given that our results distinguish three kinds of verbal stimuli (letters, word, and objects) from two kinds of non-verbal stimuli (spatial locations and gratings), we are more inclined to agree with Ruchkin and colleagues (Ruchkin, Grafman, & Krauss, 1994) in their assertion that the frontal negative wave is intimately related to a subvocal rehearsal mechanism similar to
the phonological loop proposed in Baddeley’s working memory model (Baddeley, 1985). In previous experiments, Mecklinger and colleagues themselves suggested such a relationship (Mecklinger & Meinshausen, 1998), and even their most recent work suggests that objects are rehearsed more phonologically than spatial locations (Bosch et al., 2001). While mapping scalp topographies to underlying cortex is somewhat presumptuous, the hemispheric lateralization observed between 500 and 700 ms is in agreement with functional imaging studies suggesting a role for the right hemisphere in spatial working memory and a role for the left hemisphere in verbal working memory (Jonides et al., 1993; Smith et al., 1995; Smith, Jonides, & Koepppe, 1996).

It is worth noting that increased task difficulty has been associated with increased BOLD activation in certain regions of the cortex (Barch et al., 1997) and changes in the amplitude and latency of various components of the ERP, particularly the P300 (Senkowski & Herman, 2002; Palmer, Nasman, & Wilson, 1994). Behavioral measures attest to the idea that the non-verbal tasks were more difficult than the verbal tasks in our experiment, with slower reaction time and lower accuracy for grating stimuli and lower accuracy for spatial stimuli relative the other three tasks. The accuracy-correction process previously described was performed to see whether the differences between verbal and non-verbal ERPs greatly changed when the stimulus-types were equated on accuracy, and it was found that the topographies of these differences were largely maintained. This suggests that the results obtained were not due to task difficulty, and may be due to distinctions between verbal and visual working memory. However, several alternate explanations for the observed differences in the ERP remain. For example, reaction time for grating stimuli was slower than reaction time for verbal stimuli even after accuracy-correction, and verbal and non-verbal stimulus types differ from each other on many perceptual dimensions. So while observed differences are in agreement with those found in previous neuroimaging and ERP studies, these differences cannot be confidently attributed to the difference between verbal and visual memory.

Old/New and Serial Position Effects

The ERP old/new effect is an increased positivity in the ERP observed for previously presented items relative to items presented for the first time (for review, see Johnson, 1995; Rugg, 1995). This effect begins as early as 250 ms after stimulus presentation and lasts several hundred milliseconds. Several studies have fractionated this effect into early and late parts corresponding to N400 and P300 components of the ERP, respectively. This division has been justified on experimental and theoretical grounds (Curran, 2000; Smith, 1993), as well as by evidence suggesting distinct topographies both at the scalp (Curran, 2000) and intracranially (Guillem et al., 1995). Differences in our data also defend such a dissociation. In agreement with previous findings (Curran, 2000; Smith, 1993), we found that the N400 and P300 responded differently to our experimental manipulations. Furthermore, we found that the N400 exhibited a frontal topography, while the P300 exhibited a more parietal topography. Because of this distinction, these components will be discussed separately.

N400: Familiarity and Context Integration

Dual-process theorists propose that the N400 component of the ERP is intimately tied to familiarity, which can be conceptualized as recognition in the absence of remembering
ERPs and Memory for Different Stimulus-Types

Details (Curran, 2000). However, others conceptualize the N400 as representative of the semantic integration of a stimulus with its context (Mecklinger, 1998; Rugg & Doyle, 1994). Mecklinger (1998) suggests that as a stimulus appears more often in a given context, it is better integrated and evokes a less robust (and more positive) N400. A crucial part of this framework is that semantic integration requires a “unitary code”, and because of this the context integration framework predicts that the N400 should not differentiate old and new stimuli unless those stimuli can be verbally encoded (Rugg & Doyle, 1994).

In our experiment, we found that the N400 component of the ERP old/new effect robustly differentiated old and new items for verbal stimuli, but not for non-verbal stimuli. This is consistent with other studies that failed to find an old/new difference in the ERP for non-verbal stimuli (Mecklinger, 1998, 2000; Crites et al., 2000; Bosch et al., 2001; Beisteiner et al., 1996). This provides strong support for the semantic integration hypothesis. Furthermore, this finding seems to be at odds with the dual-process theorists conceptualization of the N400, since familiarity is intuitively conceived as a feeling of knowing without the ability to verbalize the details of what one is remembering. However, making a distinction between encoding and retrieval may resolve the discrepancy between dual-process and semantic integration accounts of the N400 component of the ERP old/new effect. For example, familiarity by definition does not require the ability to verbalize the details of what is being remembered, but whether or not a familiar stimulus needed to be assigned a semantic label at the time of encoding is an entirely different question. If familiarity is dependent on some sort of semantic encoding process, then the fine line between dual-process and semantic integration accounts of the N400 blurs. Mecklinger (2000) interprets his results in a similar vein, examining the recent evidence for both dual-process (Curran, 2000; Curran & Cleary, 2003) and context integration (Mecklinger, 1998; Rugg & Doyle, 1994) accounts and concluding that his data can best be described by a combination of both. We conclude that familiarity is a process dependent on a “unitary code,” and consequently stimulus-types that cannot be encoded semantically (i.e., non-verbal stimuli) are not differentiated by the N400 component of the ERP old/new effect.

Like in previous studies of the ERP old/new effect using a Sternberg (1966) paradigm (Crittes, Devine, Lozano, & Moreno, 1998; Crittes et al., 2000), we found serial position effects in the N400 that correspond to the recency effects in our behavioral data. Third position targets ($T_3$), which are remembered faster and more accurately than targets that appeared first ($T_1$) or second ($T_2$) on the study list, also show a more positive N400. For verbal items, this fits nicely onto a continuum of memory strength, with lures showing the most negative N400, followed by $T_1$, $T_2$, and $T_3$. This is agreement with a previous study showing that the N400 becomes more positive as the strength of the memory increases (Finnigan, Humphreys, Dennis, & Geffen, 2002). Non-verbal stimuli still show a graded serial position effect, with $T_1$ showing the most negative N400 and $T_3$ showing the most positive N400, but lures show a more positive N400 than either $T_1$ or $T_2$. The fact that non-verbal stimuli show a graded serial position effect seems to be at odds with the notion that the mediation of the N400 requires semantic encoding, but Crittes et al. (2000) suggest that recency and old/new reflect two distinct processes that co-occur in time. Another possibility is that immediate recognition through familiarity does not require a semantic code, but remembering for even short durations of time requires a semantic code. This would explain the enhancement of the N400 in $T_3$ ERPs for both verbal and non-verbal stimuli. Furthermore, since $T_3$ make
the largest contributions to the N400 component of the ERP old/new effect, this explains why the effect is largely attenuated by our accuracy-correction process, which removes the majority of $T_3$ trials from analysis. If the stimulus-types were equated on behavioral measures without the selective removal of $T_3$ trials, it is possible that the effect would not be so drastically attenuated.

**P300: Recollection**

There has been substantial evidence suggesting that the P300 represents recollection, the part of recognition that dual-process theorists claim involves remembering episodic details (Smith, 1993; Duzel et al., 1997; Curran, 2000). Furthermore, in agreement with studies by Mecklinger and colleagues (Mecklinger, 1998; Bosch et al., 2001; Mecklinger, 2000), we found that the P300 component of the ERP old/new effect occurs for a range of different stimulus-types. While Mecklinger only demonstrated this for object and location stimuli, we have demonstrated that this effect occurs across a range of stimulus types, including letters, words, objects, spatial locations, and sinusoidal gratings. This suggests that recollection is a process that occurs independently of the stimulus-type remembered. Furthermore, unlike the N400 effect, the P300 component of the ERP is not sensitive to the serial position of the remembered item. This is in agreement with work by Crites and colleagues (1998). However, another study found recency effects in the P300 for non-verbal stimuli but not for verbal stimuli (Crites et al., 2000). We did not replicate this effect, finding that the P300 effect showed no recency effects for any of our stimulus-types. Our findings suggest that the P300 represents a recollective process that is independent of stimulus-type, and is capable of differentiating previously seen items from new items using a range of different stimuli. The reduction of this effect after accuracy-correction suggests that it may be sensitive to the difficulty of the task, with harder tasks eliciting smaller P300 effects. This makes sense, as harder tasks would involve items that are more difficult to recognize. This would also explain why this effect is attenuated and less widespread in our non-verbal stimuli.

**Conclusions**

We used a Sternberg (1966) paradigm and electroencephaphographic (EEG) recordings to investigate memory for different stimulus-types, including three verbal stimulus-types (letters, words, and objects) and two non-verbal stimulus-types (spatial locations, gratings). We replicated effects implicating a frontal negative slow wave in Baddeley’s (1985) phonological loop, and a parietal negative slow wave in a visuospatial rehearsal process. We also found that the early part of the ERP old/new effect (N400) was only present for verbal stimuli, but showed a recency effect for all stimulus-types, while the later P300 distinguished old and new items for the full range of stimulus-types, and never demonstrated recency effects. We conclude that familiarity, a process recently associated with the N400, requires semantic encoding for delayed but not immediate recognition, while recollection, long associated with the P300, is not dependent on a specific encoding strategy.
Acknowledgments

First and foremost, many thanks to Grace Hwang-Grodzins, Robert Sekuler, and Michael Kahana for their continued guidance and support. I would also like to thank Aaron Geller, Lynne Gauthier, Azeez Aranmolate, and Igor Korolev for their contributions to the MMS project, and Yuko Yotsumoto and Chris McLaughlin for helpful comments on previous versions of this document. Lastly, I would like to thank the other members of the Vision and Computational Memory labs for their help on countless fronts in the past year, particularly Per Sederberg and others who have contributed to the EEG toolbox.

References


ERPS AND MEMORY FOR DIFFERENT STIMULUS-TYPES


Appendix

Review of ERP Old/New Effects

Scope of Review

This review focuses largely on ERP studies of recognition memory. Particular attention is given to the differences between ERPs that are elicited by stimuli that have been previously experienced and ERPs that are elicited by new items. Research in this area fits into two broad categories: studies on the effect of item repetition during indirect tests of memory and studies on direct tests of recognition memory. Research encompassing both of these fields has focused almost exclusively on memory for words. In the interest of exploring the role of stimulus reheasability in these ERP effects, substantial time is spent reviewing ERP studies of recognition memory using non-standard stimuli, especially stimuli that resist verbal rehearsal such as spatial locations, orthographically illegal nonwords, and pictures of novel objects.

While a modified Sternberg (1966) paradigm is used in our study, little attention is afforded to ERP studies using this paradigm in this review. The reason for this is that in our experiment the number of items in the study list is not manipulated, and ERP studies using the Sternberg paradigm have focused almost exclusively on the effect of increasing list length on reaction time and the latency of the P300 component of the ERP. This is of particular interest in investigating models of short-term scanning memory. The role that ERP studies have played in the formation of such models is reviewed in depth by Kutas (1988). Several studies have investigated the difference in list-length effects between verbal and non-verbal stimuli (Osaka, 1992; Osaka & Osaka, 1983; Kotchoubey, Jordan, Grozinger, Westphal, & Kornhuber, 1996).

Event-Related Brain Potentials

The electroencephalogram (EEG) is the brain’s global electrical activity as recorded at the scalp. An event-related potential (ERP) is a time-locked voltage fluctuation at the scalp evoked by a particular mental or physical occurrence. ERPs are generally several magnitudes smaller than the ongoing EEG, but can be extracted from the EEG by averaging across several repetitions of the associated event. ERPs can be very useful for studying cognitive processes, largely because of the fine temporal resolution afforded by EEG recordings and the ability to calculate ERPs post-hoc based on subject performance measures. If different experimental conditions give rise to different ERPs, the two conditions likely evoke distinct patterns of neural activity and are likely to be functionally dissociable. Two components of the ERP that are crucial to the study of cognition are the N400 and the P300. The N400 (or N2) is a negative fluctuation in the ERP that peaks around 400 ms after stimulus presentation and has been best characterized in studies of semantic congruity (e.g., Kutas & Hillyard, 1983). The P300 is a positive fluctuation in the ERP that peaks around 300 ms after auditory stimulus presentation and around 600 ms after visual stimulus presentation, and has been implicated in a range of cognitive phenomena (Polich & Kok, 1995). This component is variably referred to as the P3, the P600, the major positive wave (MPW), or the late positive component (LPC).
ERP Repetition Effects: Indirect tests of Memory

In indirect tests of memory, the memory task is incidental to some other primary task. Typically, a list of words is presented sequentially and subjects are asked to make some decision about each word, such as in a lexical decision task. Over the course of the list, some proportion of the words are repeated one or more times. Repetition memory effects can be quantified by comparing the ERPs evoked by the first and second presentations of repeated words. One particular advantage to using this method is that, in contrast to direct tests of memory, indirect tests of memory do not confound memory effects with a certain response decision (e.g., “old” versus “new”). For example, in a lexical decision task, the first and second presentations of a single word would merit the same response. Many claim that indirect tests of memory index aspects of implicit memory. While there is some evidence for such claims in the form of studies on clinical populations with poor explicit memory (Friedman, Hamberger, Stern, & Marder, 1992; Kazmerski, Friedman, & Hewitt, 1995; Kazmeriki & Friedman, 1997; Rugg, Pearl, Walker, & Roberts, 1994), such assertions should be interpreted with caution, as recent studies have failed to replicate earlier findings that Alzheimer’s patients consistently show robust repetition effects in the ERP (Schnyer, Allen, Kaszniak, & Forster, 1999). For a thorough account of the role that studies of repetition priming have played in the concept of implicit memory, see Schacter, Chiu, and Ochsner (1993).

The sizeable literature on the effect of word repetition on the ERP has found that ERPs evoked by repeated words are more positive-going than ERPs evoked by first presentations, and that this effect peaks between 400 and 450 ms after stimulus presentation and spans both the N400 and P300 components of the ERP (Bentin, McCarthy, & Wood, 1985; Bentin & Peled, 1990; Besson, Kutas, & Van Peten, 1992; Besson & Kutas, 1993; Hamberger & Friedman, 1992; Rugg, 1985, 1987, 1990). This effect is commonly called the ERP repetition effect (Rugg, 1985). Several studies have shown that the ERP repetition effect outlasts the span of short-term memory (Karayanidis, Andrews, Ward, & McConaghy, 1991; Nagy & Rugg, 1989), and occurs for repetitions both within and across sensory modalities (Rugg, Doyle, & Melan, 1993; Rugg, Doyle, & Wells, 1995; Rugg & Nieto-Vegas, 1999). Rugg (1985) postulated that modulation of the P300 component of the ERP was largely responsible for the positive shift in repeated words, but noted with caution that the rarity of word repetitions could be responsible for the P300 shift, as has been shown using the odd-ball paradigm (Dunca-Johnson & Donchin, 1977).

Rugg and Nagy (1987) explored the ERP repetition effect for both legal and illegal nonwords. This was among the first studies investigating ERP repetition effects evoked by stimuli that resisted verbal rehearsal. Legal nonwords were constructed by swapping one or two letters of English words such that the resulting nonwords were pronounceable and orthographically legal (e.g., BLOOT, NIPER). Illegal nonwords were produced by rearranging and changing the letters of English words such that the resulting nonwords were neither pronounceable nor orthographically legal (e.g., RRWSI, TJRKDE). In the experiment, subjects were sequentially presented with a list of nonwords containing occasional repetitions. Subjects were asked to keep a silent count of the number of nonwords containing an @ character, which occurred with equal probability for legal and illegal nonwords. ERPs elicited by legal nonwords showed the standard repetition pattern that had been previously reported.
for words and nonwords, but illegal nonwords showed no such pattern. Rugg and Nagy’s findings not only suggested an important role for lexical system in the ERP repetition effect, but also quashed concerns that the effect was the result of the rarity of repetitions, since both legal and illegal nonword repetition occurred with equal probability, but only legal nonwords elicited the effect.

While a role of the P300 component could not be ruled out by their findings, Rugg and Nagy (1987) proposed a role of the N400 component in the ERP repetition effect. The N400 has been found to index stimulus expectancy and other forms of semantic priming, with unexpected or unprimed words eliciting a robust negative shift in this component (Bentin et al., 1985; Kutas & Hillyard, 1980, 1983, 1984). Rugg and Nagy suggested that previous presentation of a word primed the word such that the N400 was attenuated upon stimulus repetition, resulting in a positive shift in the ERP. While Rugg and Nagy’s work was an excellent first step toward understanding the dynamics of ERP repetition effects for non-verbal stimuli, their illegal nonwords could still be considered verbal, as they could be verbally labeled as a list of letters.

Rugg and Doyle (1994) argued that the N400 component represents the degree to which a word can be integrated with its context. According to what has been called the context integration hypothesis, contextual integration is facilitated when a word re-occurs one or more times in the same context. Rugg and Doyle interpreted Rugg and Nagy’s (1987) results within this framework, suggesting that the amplitude of the N400 decreases with contextual integration, which results in a more positive-going ERP. Furthermore, Rugg and Doyle claimed that context integration is dependent on certain stimulus properties, particularly the ability of the stimulus to be represented as a semantic construct that they refer to as a “unitary code.” In the case of Rugg and Nagy’s study, while legal nonwords can be represented as unitary semantic codes corresponding to their pronunciations, illegal nonwords can be represented minimally as a list of letters.

Rugg and Doyle (1994) describe an experiment (further discussed in Thomas, 1992; Jordan & Thomas, 1999) that investigated ERP repetition effects for meaningful and non-meaningful pictures. The meaningful objects consisted of randomly selected photographs of scenes and objects. The non-meaningful objects were similarly selected photographs that were distorted by removing high spatial frequencies, and were therefore highly ambiguous. While the two groups of pictures were similar in many physical properties, including color, size, and brightness, they differed markedly in their spatial frequency content. Rugg and Doyle described robust repetition effects in the ERP evoked by the meaningful pictures, but no such effects were apparent in the ERPs evoked by the non-meaningful pictures. They emphasize that manipulating meaningfulness in pictures has a similar effect on ERP repetition effects to manipulating orthographic legality in words. Similarly to their interpretation of Rugg and Nagy’s (1987) results, they claim that the ERP repetition effect results from an attenuation of the N400 component, and that meaningful pictures show a marked repetition effect because they can be assigned a unitary code, while non-meaningful pictures are more difficult to label. It is important to interpret these results with caution, as an alternative hypothesis exists whereby the differences between the meaningful and non-meaningful pictures can be explained solely on the basis of differences in spatial frequency content. It is possible, for example, that repetition effects represent a priming mechanism that is wholly dependent on high-frequency information. However, it is telling that two studies using very
different stimuli both find ERP repetition effects for the more meaningful stimuli and no repetition effects for less meaningful, harder-to-label stimuli.

Rugg, Soardi, and Doyle (1995) used an indirect memory task to investigate the repetition effects for line drawings of novel 3-dimensional objects. In the first of the two experiments they described, Rugg et al. used a primary task in which subjects were asked to respond to impossible objects (i.e., objects that could not exist in 3-dimensional space) against a background of possible objects, some proportion of which repeated immediately after the first presentation. In the second experiment, the primary task was to respond whenever the object contained a pair of parallel lines. Surprisingly, results from the first experiment yielded negative-going repetition effects starting at approximately 300 ms and lasting the duration of the ERP, in striking contrast to the positive-going ERP repetition effects typically elicited by words. Distinct negative-going repetition effects in the earliest part of the ERP had previously been reported by Otten, Rugg, and Doyle (1993) for stimuli in unattended areas of the visual field, but this negativity gave way to a positive-going repetition effect in later parts of the ERP. Rugg et al. argued that the effect could not be explained in terms of differential eye movements to repetitions and was likely to be a genuine negative ERP repetition effect. In agreement with previous findings by Rugg and Nagy (1987) and Rugg and Doyle (1994), experiment 2 showed no ERP repetition effects for novel 3-dimensional stimuli. This is interesting in light of the observation that it was trivial for subjects to recognize the repetitions both in this study and in the studies using frequency-filtered photographs (Jordon & Thomas, 1999; Thomas, 1992). Van Petten and Senkfor (1996) corroborate the findings from the second experiment by finding positive-going ERP repetition effects in study lists of words, but not of novel visual patterns. Similarly, Zhang, Begleiter, Projesz, and Litke (1997), found ERP repetition effects for meaningful words and images but not for meaningless scrambles of these images and words. However, the differential findings between Rugg, Soardi, and Doyle’s (1995) two experiments uncover some of the flaws of indirect memory tests, particularly the effect of the primary task on memory effects.

In contrast to the results found by other researchers using novel visual stimuli, a study using novel 3-dimensional object stimuli found positive-going ERP repetition effects for possible but not impossible novel objects (Penney et al., 2000). The principal difference is that Penney et al. emphasized to their subjects the importance of encoding the structural aspects of the stimuli, whereas the subjects in the other studies only encoded the stimuli to the degree necessary to perform the primary task. It is entirely possible, for example, that subjects under the additional pressure to encode the stimuli found ways of labeling each stimulus with what Rugg and Doyle would call a “unitary code.”

Overall, studies of the ERP repetition effect have found robust, positive-going effects in the ERP for words (Rugg, 1985), legal nonwords (Rugg, 1987), and meaningful pictures (Zhang et al., 1997). However, with one exception (Penney et al., 2000), studies of less meaningful, non-verbal stimuli such as novel images (Jordon & Thomas, 1999; Van Petten & Senkfor, 1996; Rugg et al., 1995; Thomas, 1992; Zhang et al., 1997) and illegal nonwords (Rugg & Nagy, 1987) have found no positive shift in the ERP after repetition (Rugg & Doyle, 1994). These results are interpreted in terms of the context integration hypothesis, and point to the importance of a “unitary code” in the ERP repetition effect.
In direct tests of recognition memory, the subjects primary task is to indicate whether the stimulus has been previously experienced ("old") or not ("new"). Thus, as with indirect tests of memory, the comparison of interest is between items that have been presented only once and items that have been presented twice. There are two major types of direct tests. The first and most popular type of direct test has a study-test format, with a list of study items followed by a list of test items. Subjects are asked to indicate which test items were on the study list and which were not. A second type of direct test is a continuous recognition task, in which only one list is presented and subjects are asked whether each item has been previously presented or has not. Among the advantages of the direct tests of memory is that, unlike indirect tests, subject performance measures can be collected in addition to electrophysiological measures. However, memory effects in direct tests of memory are often confounded by response choice because remembered items warrant one response ("old") and non-remembered items elicit a different response ("new").

In agreement with results of ERP studies using indirect tests of memory, ERP studies using direct tests generally find more positive-going ERPs throughout the N400 and P300 components for the old words compared to the new words (for review, see Johnson, 1995; Rugg, 1995). This difference is commonly called the ERP old/new effect, and has been theoretically divided into early and late components corresponding to the N400 and P300, respectively (Smith & Halgren, 1989). Rugg, Brovedani, and Doyle (1992) connected studies using indirect and direct tests of memory by showing that the memory effect in the ERP occurred independently of response choice. Subjects performed a continuous recognition memory test within each of five test blocks. However, half of the items in the second through fifth blocks were repeated items from previous blocks. Rugg et al. (1992) found an ERP old/new effect for items repeated within blocks, and a practically identical ERP repetition effect for items repeated across blocks. While this suggests that similar phenomena may underlie the ERP repetition effect found in indirect tests of memory and the ERP old/new effect found in direct tests of memory, the relationship between these two effects remains largely unclear.

**Dual-Process Theory**

Most of the research using direct tests of memory has focused on dual process theories of recognition memory. According to dual process theorists, dissociations observed in recognition memory tasks can be explained by two separable memory processes: recollection and familiarity (for review, see Yonelinas, 2002). Dual process theorists generally describe recollection as a process involving the recognition of specific details, including episodic information, while they describe familiarity as an assessment of the global similarity between a probe item and the contents of memory without taking detailed source information into account. The classic illustration of this dissociation is the experience of being familiar with a person’s face without being able to recollect the person’s name (Mandler, 1991).

Early studies interpreting the ERP old/new effect within a dual process framework yielded a hodgepodge of relatively unsupported claims. Based on their studies using left- and right-hemisphere damaged patients, Smith and Halgren (1989) claimed that the N400 and P300 were indices of a recollective process, and suggested that differences in familiarity...
were not reflected in the ERP. However, when Potter, Pickles, Roberts, and Rugg (1992) used scopolamine to impair subject performance on a continuous recognition task, they found that the drug had no effect on the early ERP old/new effect and actually enhanced the late ERP old/new effect. Based on their findings and the presumption that scopolamine selectively impairs recollection, Potter et al. (1992) claimed that these components could not represent a recollective process, and proposed that the two components represented a familiarity-based process. The dispute remained divided for many years, and further claims were made that the ERP old/new effect indexed the process of familiarity (e.g., Rugg & Doyle, 1992, 1994) or the process of recollection (e.g., Gardiner & Java, 1990; Paller & Kutas, 1992).

One variation on the study-test procedure that is of particular interest to dual-process theorists is the remember/know procedure (Tulving, 1985). The remember/know procedure is identical to the study-test procedure in all regards except that when subjects recognize a stimulus they must indicate whether they simply know that the item was previously studied or whether they could remember the details of its original presentation. When interpreted within a dual-process framework, the remember/know procedure yields responses corresponding to both recollection ("remember") and familiarity ("know"). Smith (1993) used the remember/know paradigm in order to understand functional significance of the ERP old/new effect. Smith found that the early ERP old/new effect (N400) did not differ significantly between remember and know judgments, but that the late ERP old/new effect (P300) was considerably larger for remember judgments compared to know judgments. Smith concluded that the P300 represented a recollective process, while the N400 may be an index of repetition priming.

Wilding and colleagues (Wilding et al., 1995; Wilding & Rugg, 1996, 1997) performed a series of experiments using a study-test procedure with the additional requirement of a source judgment along with the standard old/new judgment. For example, after the initial old/new judgment, subjects were asked whether the test word was presented auditorily or visually (Wilding et al., 1995) or whether the word was spoken in a male or female voice (Wilding & Rugg, 1996). Experiments using this procedure have been found to be comparable to those using the remember/know procedure, with correct source attribution eliciting similar ERP effects to remember judgments (Rugg et al., 1998). Wilding and colleagues found that the ERP old/new effect was significantly larger when a correct source judgment was made when compared to when an incorrect source judgment was made, but that these two effects did not differ is scalp distribution. Although their findings supported the idea that the ERP old/new effect represented a recollective process, Wilding and colleagues questioned the validity of dual process theories. Because source recognition caused a quantitative rather than qualitative change in the ERP, they suggested that familiarity and recollection were not truly independent processes, and could be better conceived as existing on a continuum from weak to strong memory strength.

However, Duzel, Yonelinas, Mangun, Heinze, and Tulving (1997) used a remember/know paradigm and found qualitative differences in the ERP old/new effect between remember and know judgments, with know judgments being associated with a fluctuation in the early N400 component, and remember judgments being associated with a change in the later P300 component. Curran and colleagues have corroborated these finding in a series of experiments using the remember/know procedure (Curran, 2004) and a variation
on the study-test procedure in which subjects are expected to differentiate between studied stimuli and new stimuli that are sometimes very similar to the studied stimuli (Curran, 2000; Curran & Cleary, 2003). In one experiment, some of the new stimuli were words from the study list with changed plurality (Curran, 2000), and in another experiment they were mirror images of meaningful pictures from the study list (Curran & Cleary, 2003). In both cases, Curran and colleagues found that remembrance of the specific details involved in identifying old items involved a positive fluctuation of the P300 component, but similar new items identified as old had no effect on the P300. This is consistent with the P300 being an index of recollection, a process that involves the recognition of details (i.e., plurality and orientation). The N400 (dubbed the “FN400” by Curran) was more positive for both old items and similar new items relative to other new items, possibly because both studied items and new items similar to studied items would evoke feelings of familiarity. The work by Curran and colleagues (Curran, 2000; Curran & Cleary, 2003; Curran, 2004) and Duzel et al. (1997) suggest that the early part of the ERP old/new effect (N400) may index familiarity and the later part of the ERP old/new effect (P300) may index a recollective process, which is consistent with the presumption that familiarity precedes recollection in time (Yonelinas, 2002).

The Effect of Stimulus-Type.

There have been several recent studies using direct tests of recognition memory with non-verbal stimuli, such as novel shapes or visual objects that are difficult to label. For example, Crites, Delgady, Devine, and Lozano (2000) used a Sternberg (1966) short-term scanning paradigm to investigate the differences in the ERP old/new effect between visual images that were easy to verbally label (EL task) versus images that were difficult to label (DL task). Images in the EL task were pictures of common objects, and images in the DL task consisted of various ambiguous color and form patterns. Interestingly, Crites et al. found recency effects in the N400 component of the ERP for both the EL task and the DL task, with more recently presented old stimuli having more positive-going ERPs than more distally presented stimuli. This confirmed and expanded the findings of a previous study that used only verbal stimuli (Crites et al., 1998). Consistent with previous findings using the ERP old/new and repetition effects, Crites et al. found that ERPs evoked by old probes were more positive-going than ERPs to new probes, but that this effect only occurred during the EL task. This is consistent with findings that the ERP repetition effect does not occur for non-verbal stimuli (e.g., Rugg & Doyle, 1994; Rugg & Nagy, 1987). Furthermore, Beisteiner et al. (1996) had similar findings using a continuous recognition task for non-words and geometric figures, with non-words showing an ERP old/new effect but no effect being apparent for the figures.

Conversely, one study in which subjects were trained to categorize novel visual blobs into groups found ERP old/new effects across both the N400 and P300 components of the ERP during a study-test procedure for these stimuli (Curran et al., 2002). The training that subjects in this experiment received may have allowed them to verbally label the blobs based on group membership, and may explain the presence of ERP old/new effects to these supposedly non-verbal stimuli despite contradicting evidence from other studies.

Mecklinger and colleagues performed a series of studies using a study-test procedure to compare memory for common objects with memory for object locations (Bosch et al.,
ERPS AND MEMORY FOR DIFFERENT STIMULUS-TYPES

2001; Mecklinger & Muller, 1996; Mecklinger, 1998; Mecklinger & Bosch, 1999; Mecklinger & Meinshausen, 1998). Although intended to be non-verbal, the object stimuli used in these studies were typically easily verbally labeled and consisted of common shapes and line drawings of everyday objects, with the exception of one study in which letters of the Klingon alphabet were substituted for the shapes (Mecklinger & Bosch, 1999). Furthermore, the locations fit into one of between eight and 16 positions in a 2-dimensional (Mecklinger & Muller, 1996; Mecklinger, 1998; Mecklinger & Meinshausen, 1998; Mecklinger & Bosch, 1999) or 3-dimensional (Bosch et al., 2001) spatial matrix. In the object condition, subjects were expected to remember the objects but not their locations, and in the spatial condition the subjects were asked to remember the objects locations but not the objects themselves. Importantly, the two conditions were identical except for the instructions given. A simple articulatory suppression task was used in an attempt to suppress verbal rehearsal, but the effectiveness of said task is questionable. Mecklinger and colleagues found qualitatively different ERP old/new effects in the object and spatial conditions. In the object condition, Mecklinger and colleagues found a frontally-centered, N400 effect similar to that found by Curran (2000), whereas they found a parietal P300 effect in the spatial condition. Mecklinger (2000) analyzed these findings with respect to dual-process theories and the concept of a “unitary code.” Mecklinger suggested that no N400 old/new effect was present in the spatial task because the object locations could not be assigned the “unitary code” necessary for familiarity-based recognition, and that a more detail-oriented recollective process was necessary for successful recognition of these locations. However, both recollective and familiarity-based processes could be used to recognize the objects themselves. This fits with Curran’s (2000) concept that the N400 and P300 are indices of familiarity and recollection, respectively, and that these two processes may be drawn on to different degrees in different tasks.

Studies using direct tests of memory have focused largely on dual-process accounts of recognition memory. Consistently, it has been found that items recognized as old elicit more positive-going ERPs than new items. Recent evidence (Duzel et al., 1997; Curran, 2000; Curran & Cleary, 2003; Curran, 2004) has suggested that the early component responsible for this effect (N400) may index a familiarity-based process and the later component (P300) may index a recollective process. The few studies using non-verbal stimuli in direct tests of memory have suggested that the old/new effect is absent when the stimuli used resist verbal rehearsal, but some contention remains.