Stages of Processing in Associative Recognition: Evidence from Behavior, EEG, and Classification

Jelmer P. Borst*, Darryl W. Schneider*, Matthew M. Walsh*, and John R. Anderson

Abstract

■ In this study, we investigated the stages of information processing in associative recognition. We recorded EEG data while participants performed an associative recognition task that involved manipulations of word length, associative fan, and probe type, which were hypothesized to affect the perceptual encoding, retrieval, and decision stages of the recognition task, respectively. Analyses of the behavioral and EEG data, supplemented with classification of the EEG data using machine-learning techniques, provided evidence that generally supported the sequence of stages assumed by a computational model developed in the Adaptive Control of Thought-Rational cognitive architecture. However, the results suggested a more complex relationship between memory retrieval and decision-making than assumed by the model. Implications of the results for modeling associative recognition are discussed. The study illustrates how a classifier approach, in combination with focused manipulations, can be used to investigate the timing of processing stages. ■

INTRODUCTION

A longstanding interest of cognitive psychologists and neuroscientists has been the development of methods to identify different stages of human information processing. For example, the subtractive method of Donders (1969) involves comparing tasks that are hypothesized to share all but one processing stage, then making inferences about the duration of that stage based on differences in RT between tasks. The additive factors method of Sternberg (1969) involves studying the effects of various experimental manipulations on RT, with additive effects being diagnostic of separate stages. Such methods allow for certain inferences about processing stages from behavioral data, but they are fundamentally limited in that RT is a coarse-grained measure that reflects the cumulative duration of all stages.

EEG is a method that can provide a more fine-grained view of different stages of information processing. The temporal precision afforded by recording neurally generated electrical signals on a millisecond basis allows researchers to determine whether, when, and for how long an experimental manipulation modulates the neural activity underlying cognitive processes, even in cases where the manipulation does not produce an effect on overall RT (Coles, 1988). Such knowledge can be useful for evaluating computational models of cognition that assume a particular composition of processing stages for performing a task. The purpose of this study was to use EEG data to evaluate the hypothesized information processing stages in a computational model for the task of associative recognition.

Associative Recognition

Associative recognition involves judging whether two items were previously experienced together. For example, the task used in this study involved determining whether a probe stimulus consisted of two words that were studied together (targets) or separately (re-paired foils). Successful discrimination required remembering not only that the words were studied (item information) but also how the words were paired during study (associative information).

Among the many computational models of associative recognition is a process model based on the Adaptive Control of Thought-Rational (ACT-R) theory, which is an integrated architecture for modeling cognition (Anderson, 2007). The ACT-R model assumes associative recognition is accomplished by a sequence of four processing stages. The first stage involves encoding a perceptual representation of the probe presented for recognition. Perceptual encoding is sensitive to probe features such as the size of a visually presented word pair. The second stage involves using the encoded representation to retrieve a studied word pair from memory. Retrieval is sensitive to

Carnegie Mellon University
*These authors contributed equally to this work and order of authorship is alphabetical.

© 2013 Massachusetts Institute of Technology

Journal of Cognitive Neuroscience 25:12, pp. 2151–2166
doi:10.1162/jocn_a_00457

the associative fan of the words used to access memory, as discussed below. The third stage involves comparing the retrieved word pair with the probe word pair and deciding whether they match. The decision process is sensitive to the type of probe, with targets resulting in matches and re-paired foils resulting in mismatches. The fourth stage involves executing the appropriate response based on the outcome of the matching process (e.g., a keypress indicating "yes" or "no" with respect to whether the probe was studied). Responding is sensitive to motor features such as the hand and the finger used to make a keypress response.

The four stages in the ACT-R model are assumed to occur in serial order, which implies that manipulations affecting each stage should be manifest at different times during task performance. For example, the effect of a perceptual encoding manipulation should be evident before the effect of a retrieval manipulation. In this study, we used detailed temporal data from EEG to provide converging evidence regarding the putative stages of associative recognition represented in the ACT-R model. We recorded EEG data while participants performed an associative recognition task that involved three manipulations, each of which was intended to tap one or more of the first three processing stages in the model.

Word Length

The first manipulation pertained to the perceptual encoding stage and involved presenting probe word pairs that consisted of either short (four- or five-letter) or long (seven- or eight-letter) words. EEG studies of word length have yielded conflicting results, in part because word length and word frequency are usually confounded—shorter words occur more frequently in natural language corpora (Zipf, 1935). However, studies that control for word frequency provide a more consistent picture. Long words produce a greater positivity over the occipital region beginning around 85 msec (Hauk, Pulvermüller, Ford, Marslen-Wilson, & Davis, 2009; Hauk, Davis, Ford, Pulvermüller, & Marslen-Wilson, 2006; Hauk & Pulvermüller, 2004), followed by a broad positivity over the frontal region beginning around 300 msec (Van Petten & Kutas, 1990). The ERP component underlying the early word length effect, called the P1, is thought to reflect low-level perceptual analysis of visual stimuli (Dien, 2009). The ERP component underlying the later frontal positivity, however, is less understood. On the basis of these findings, we predicted that long words would produce more positive-going waveforms over the parietal scalp around 85 msec and over the frontal scalp around 300 msec.

Associative Fan

The second manipulation pertained to the retrieval stage and involved having probe words with different associative fans, which refers to the number of episodic associations that a word has with other words in memory. A well-established phenomenon in associative recognition is the fan effect, which is the finding that RT becomes longer as fan increases (e.g., Pirolli & Anderson, 1985; Anderson, 1974; for reviews, see Anderson, 2007; Anderson & Reder, 1999). The ACT-R model explains the fan effect as a form of associative interference, such that the activation that a probe provides for items in memory decreases as its fan increases. Retrieval time is inversely related to activation, resulting in longer RTs for higher fan probes (Schneider & Anderson, 2012; Anderson & Reder, 1999).

Besides expecting a fan effect on RT, we were interested in whether fan would be reflected in the EEG signal. Using a modified fan paradigm, Heil, Rösler, and Hennighausen (1996, 1997) discovered negative slow potentials that accompanied the reactivation of studied information. The amplitude of these potentials increased with the amount of material retrieved (i.e., the number of items associated with the probe). The topographic distribution of this effect was specific to the modality of the studied material: words evoked a maximum negativity over left-frontal sites, spatial configurations evoked a maximum negativity over parietal sites, and faces evoked a maximum negativity over left-central sites (Khader et al., 2007; Khader, Heil, & Rösler, 2005; Heil et al., 1997). Nyhus and Curran (2009) studied recognition memory using a task in which the associative fan of the font of presented words was manipulated. They found that associative fan modulated the EEG signal at left frontal and parietal sites. Finally, fMRI studies of the fan effect have reported greater activity in left pFC as participants retrieve high-fan items (Danker, Gunn, & Anderson, 2008; Sohn, Goode, Stenger, Carter, & Anderson, 2003). On the basis of these findings, we predicted that differences in associative fan for word stimuli would produce modulation of the EEG waveform over the frontal and parietal scalp, with the strongest effects appearing in the left hemisphere.

Probe Type

The third manipulation pertained to the retrieval and decision stages and involved having probes that either matched or mismatched studied word pairs (targets and re-paired foils, respectively). In the case of a target, the ACT-R model assumes that the retrieval stage results in the retrieval of a matching word pair because both probe words provide activation for the same studied word pair in memory. In the case of a re-paired foil, the model assumes that retrieval produces a mismatching word pair because the probe words provide partial activation for different studied word pairs in memory. According to the model, the lower activation for retrieval and the resulting mismatch would prolong RT for re-paired foils relative to targets.

Besides expecting a probe effect on RT, we were interested in whether probe type would be reflected in the

EEG signal. Previous EEG research supports a dual-process theory of recognition (for a review, see Rugg & Curran, 2007; see also Diana, Reder, Arndt, & Park, 2006; Yonelinas, 2002) involving two qualitatively distinct processes: familiarity and recollection. Familiarity is a fast and context-free process that provides an initial sense of whether the probe was studied. Recollection is a slow and effortful process that entails retrieving details about studied pairs. It is generally agreed that associative recognition involves recollection (e.g., Yonelinas, 2002; but see Speer & Curran, 2007, for a discussion on possible involvement of familiarity processes), which is akin to the retrieval process in the ACT-R model.

In EEG studies of recognition memory, recollection is associated with more positive going waveforms for targets than for foils over the left-parietal scalp beginning 450 msec after stimulus presentation, a finding known as the parietal old/new effect (e.g., Curran, 2000; Düzel, Yonelinas, Mangun, Heinze, & Tulving, 1997). It has been suggested that the effect is sensitive to the amount of information recollected (Vilberg & Rugg, 2009; Vilberg, Moosavi, & Rugg, 2006; Wilding, 2000), although the relevant evidence comes from judgments of recollected details rather than manipulations of associative fan. Likewise, Nyhus and Curran (2009) found that the parietal old/new effect interacted with associative fan. Given that the retrieval stage of the ACT-R model is sensitive to probe type (targets are retrieved faster than re-paired foils), we predicted a parietal old/new effect during the retrieval stage, which is modulated by fan. Furthermore, because the decision stage of the ACT-R model involves matching based on the outcome of retrieval, the parietal old/new effect was predicted to extend into the decision stage.

In summary, we expected to find an effect of word length during the encoding stage, effects of fan and probe type during the retrieval stage, and an effect of probe type during the decision stage. The manner in which the effects of these manipulations appear in the EEG data can provide evidence for or against the stages assumed by the ACT-R model. For example, if the onsets of EEG modulations related to word length, fan, and probe type occur in that order at reasonable times, then it would represent evidence in support of the model. However, if the effects occur in a different order or the nature of an effect changes over time (e.g., the effect reverses), then it would represent evidence against the model. In that case, the observed EEG pattern may suggest revisions to the model that would make it more compatible with the data.

Classification Analysis

The EEG data can also be used in a model-free way to explore the processing stages involved in associative recognition. Recent advances in machine learning classification allow one to use neural data to identify different stages of cognitive processing. For example, machine learning classifiers have been used to analyze fMRI data (multivoxel pattern analysis or mind reading: e.g., Pereira, Mitchell, & Botvinick, 2009; Haynes & Rees, 2006; Norman, Polyn, Detre, & Haxby, 2006), EEG data (brain–computer interfacing or decoding: e.g., Das, Giesbrecht, & Eckstein, 2010; Peters, Pfurtscheller, & Flyvbjerg, 1998), magnetoencephalography (MEG) data (e.g., Parra et al., 2002), and simultaneous EEG and MEG recordings (Chan, Halgren, Marinkovic, & Cash, 2011). Instead of using univariate methods to investigate which voxels or electrodes respond to the conditions of the experiment, as most classical analysis methods do, classifier approaches inspect how information is spatially and/or temporally represented by a combination of voxel or electrode values. For instance, classifiers have been used to investigate the representation of semantic categories (fMRI: Mitchell et al., 2008; MEG: Sudre et al., 2012), stages in solving algebra problems (fMRI: Anderson, Betts, Ferris, & Fin.am, 2011), distinctions between faces and cars (EEG: Philiastides & Sajda, 2006), and states in a memory game (fMRI: Anderson, Fincham, Schneider, & Yang, 2012).

To analyze the processing stages in our study, we trained a classifier with EEG data to identify different experimental conditions. We trained and tested the classifier on 50-msec windows between stimulus onset and response generation to see when information related to the conditions became available. The logic of the approach is that if the classifier can distinguish between two conditions during a certain time period (e.g., short and long words from 100 to 200 msec), then one can conclude that there is information in the EEG data that distinguishes between those conditions at that time. The time course and spatial representation of the classification can then be used to make inferences about the processing stages involved in performing the task.

METHODS

The experiment consisted of two phases: a training phase and a test phase. In the training phase, participants learned 32 word pairs by completing a cued recall task. In the subsequent test phase—during which EEG data were collected—participants completed an associative recognition task in which they distinguished targets (trained word pairs) from re-paired foils (alternative pairings of trained words) and new foils (pairs of novel words not presented during training) In addition to probe type, we manipulated word length (words were either short [four or five letters] or long [seven or eight letters]) and associative fan (words had either one or two associates).

Participants

Twenty individuals from the Carnegie Mellon University community each participated in a single 3-hr session for monetary compensation (9 men and 11 women, ages ranging from 18 to 40 years with a mean age of 26 years).

All were right-handed, and none reported a history of neurological impairment.

Materials

Word pairs were constructed from a pool of 464 words selected from the MRC Psycholinguistic Database (Coltheart, 1981). The words were nouns with word frequency between 2 and 100 occurrences per million and a minimum imageability rating of 300. Half of the words were four or five letters and composed the short word list, which had a mean word frequency of 24.3 occurrences per million (SD = 22.1), mean imageability rating of 539.3 (SD = 55.3), and mean word length of 4.5 letters (SD = 0.5). The other half of the words were seven or eight letters and composed the long word list, which had a mean word frequency of 24.4 occurrences per million (SD = 23.4), mean imageability rating of 505.6 (SD = 81.6), and mean word length of 7.2 letters (SD = 0.4). The 232 words of each length were divided randomly into two lists—a 24-word study list and a 208-word new foil list—such that the lists were matched on word frequency, imageability, and word length according to t tests (all ps > .1). Word frequency was also matched across the corresponding lists of each length, thereby avoiding the natural confound between word frequency and length. Study lists were also constrained such that each word started with a unique three-letter sequence.

The lists were used to create three sets of probes: targets, re-paired foils, and new foils. A set of 32 target word pairs was constructed from the study lists such that there were eight word pairs for each combination of length (short or long) and fan (1 or 2). Both words in short pairs were four or five letters, and both words in long pairs were seven or eight letters. Each word in a Fan 1 pair appeared only in that pair, whereas each word in a Fan 2 pair appeared in two pairs. A corresponding set of 32 re-paired foil word pairs was constructed in a similar manner by recombining words from different target pairs of the appropriate length and fan. A set of 208 new foil word pairs was constructed from the new foil lists such that there were 104 word pairs for each length (all Fan 1). Thus, there were 10 conditions defined by the probes, reflecting the eight conditions from a 2 (probe: target or re-paired foil) × 2 (length: short or long) × 2 (fan: 1 or 2) design and two additional conditions represented by short and long new foils.[1] The randomization of words and their assignment to conditions were unique for each participant.

Procedure

The experiment began with a training phase in which participants learned the target word pairs. The training phase started with each target word pair presented onscreen (one word above the other) for 5000 msec and followed by a 500-msec blank screen. Participants were instructed to read each pair and make an initial effort to memorize it. Following target presentation, participants completed a cued recall task designed to help them learn the target word pairs. On each trial, they were presented with a randomly selected target word and had to recall the word(s) paired with it (two-word responses were required for Fan 2 words). The self-paced responses were typed and feedback (in the form of the correct response) was provided for 2500 msec following errors. If a target word elicited an error, it was presented again after all other target words had been presented. A block of trials concluded when all 48 target words had elicited correct responses. Participants completed a total of three blocks of cued recall.

After the training phase, participants entered the EEG recording chamber and completed the test phase. Each trial began with a centrally presented fixation cross for a variable duration (sampled from a uniform distribution ranging from 400 to 600 msec). Following fixation, a probe word pair appeared onscreen (one word above the other) until the participant responded with a keypress to indicate whether the probe had been studied during the training phase. The probe was either a target, re-paired foil, or new foil. Targets required "yes" responses (indicated by pressing the J key with the right index finger) and foils required "no" responses (indicated by pressing the K key with the right middle finger). Participants made all responses with the right hand to avoid confounding probe effects with bilateral motor potentials (e.g., the lateralized readiness potential; Smulders & Miller, 2013). Participants were instructed to respond quickly and accurately. Following the response, accuracy feedback was displayed for 1000 msec, after which a blank screen appeared for 500 msec before the next trial began. Participants completed a total of 13 blocks with 80 trials per block. All 10 conditions occurred equally often in random order in each block, resulting in 104 trials per condition during the test phase. All targets and re-paired foils appeared once per block and were thus presented 13 times during the test phase. Each new foil appeared only once during the test phase.

EEG Recording and Analysis

Participants sat in an electromagnetically shielded chamber. Stimuli appeared on a CRT monitor placed behind radio frequency shielded glass and set 60 cm from participants. The EEG was recorded from 32 Ag-AgCl sintered electrodes (10–20 system). Electrodes were also placed on the right and left mastoids. The right mastoid served as the reference electrode, and scalp recordings were algebraically re-referenced off-line to the average of the right and left mastoids. The vertical EOG was recorded as the potential between electrodes placed above and below the left eye, and the horizontal EOG was recorded as the potential between electrodes placed at the external canthi. The EEG and EOG signals were amplified by a Neuroscan bioamplification system with a bandpass of 0.1–70.0 Hz and were digitized at 250 Hz. Electrode impedances were kept below 5 kΩ.

The EEG recording was decomposed into independent components using the EEGLAB infomax algorithm (Delorme & Makeig, 2004). Components associated with eye blinks were visually identified and projected out of the EEG recording. Stimulus-locked epochs of 1200 msec (including a 200-msec baseline) were then extracted from the continuous recording and baseline-corrected using the prestimulus interval. Response-locked epochs of 1200 msec (including a 600-msec pre-response period) were also extracted from the continuous recording. Response-locked analyses allowed us to examine late-occurring effects of fan and probe while controlling for variability in the time to respond to the different probes. Because we were interested in decision-related effects rather than response-related effects, we corrected response-locked epochs using the 200-msec prestimulus baseline (Luck, 2005). Stimulus- and response-locked epochs containing voltages above +75 μV or below −75 μV were excluded from further analysis. Furthermore, trials with incorrect responses and trials with RTs more than 3 *SD*s longer than the mean correct RT for a given condition and participant were excluded from analysis.

To identify time periods of interest, we computed average voltages over adjacent 50-msec windows at each electrode, both for the stimulus- and response-locked ERPs. We then performed a 2 (Length: short or long) × 2 (Fan: 1 or 2) × 2 (Probe: target or re-paired foil) repeated-measures ANOVA at each electrode and for each time window (32 electrodes × 24 windows). We identified windows where electrodes showed significant main effects following false discovery rate (FDR) correction for multiple comparisons (Genovese, Lazar, & Nichols, 2002), using an FDR of 0.01.[2] Although it was not a necessary outcome of this analysis, significant electrodes within each of the temporally defined windows were contiguous and showed consistent effects. Thus, we used the FDR analysis to construct ROIs involving the subsets of electrodes showing significant effects within each of the temporally defined windows. We used repeated-measures ANOVAs to examine the effects of length, fan, and probe on ERPs recorded over the FDR-defined ROIs.

Classifier

To classify the data, we followed the methodology outlined by Sudre et al. (2012), who used a classifier in combination with MEG data to investigate when perceptual and semantic features of nouns were reflected in neural activity. We first discuss how we preprocessed the data for the classifier, and then we describe the algorithm and the training/testing methodology.

Classifier Preprocessing

Because the classifier algorithm requires that trials be the same length, we stimulus- and response-locked the data, which enabled us to capture effects occurring at both the start and the end of the trials in a synchronized manner. Trials were adjusted to be 1200 msec in length, with the first 500 msec constructed from stimulus-locked data and the remaining 700 msec constructed by resampling the period of each trial that occurred after 500 msec but before the response. In other words, the portion of each trial occurring after 500 msec was "shrunk" or "stretched" to a duration of 700 msec. This approach preserved peaks in the EEG waveforms that were present at the start and the end of a trial. This type of event-locking procedure has also been used to align individual trials of varying durations in fMRI experiments (Anderson et al., 2008).

After stimulus- and response-locking the data, we created classifier examples by averaging over the 13 presentations of each word pair for each participant. This resulted in eight examples for each of the eight conditions formed by the factorial combination of fan, length, and probe, creating 64 examples in total. Classifier epochs of 1400 msec (including a 200-msec prestimulus interval) were extracted from the 32 channels. Before creating epochs, we applied a 0.5–30 Hz band-pass filter to attenuate low- and high-frequency noise. The data were recorded at 250 Hz, yielding 351 × 32 = 11,232 features per example, where each feature is a time point at a certain channel. From the examples and features, we created a 64 × 11,232 matrix $\mathbf{X}$ for each participant.

The final step in preprocessing the data was normalizing each channel in each row in $\mathbf{X}$ to a mean of 0 and a standard deviation of 1. Each channel was normalized separately to prevent channels with higher amplitudes from disproportionately influencing the classifier results. Normalizing also ensured that different examples received equal weight. In addition to matrix $\mathbf{X}$, we also created a 64 × 3 matrix $\mathbf{Y}$ that contained the labels for the examples. The columns in $\mathbf{Y}$ coded fan (Fan 1 = −1 and Fan 2 = 1), length (short = −1 and long = 1), and probe (re-paired foil = −1 and target = 1).

Classifier Algorithm

To classify the data, we used a machine-learning technique called ridge regression (e.g., Hastie, Tibshirani, & Friedman, 2001). In essence, ridge regression is a linear regression method that can handle situations with many correlated predictors, as is the case here. Ridge regression learns a mapping $\hat{\mathbf{W}}$ between the $n \times p$ example matrix $\mathbf{X}$ and the labels $\mathbf{Y}$:

$$\hat{\mathbf{W}} = \left(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}_p\right)^{-1}\mathbf{X}^T\mathbf{Y} \quad (1)$$

$\mathbf{I}_p$ is the $p \times p$ identity matrix and λ is a complexity parameter that penalizes large coefficients $\hat{\mathbf{W}}$. Equation 1 requires the inversion of a $p \times p$ matrix. Equation 1 can be rewritten in its dual form:

$$\hat{\mathbf{W}} = \mathbf{X}^T\left(\mathbf{X}\mathbf{X}^T + \lambda\mathbf{I}_n\right)^{-1}\mathbf{Y} \quad (2)$$

in which $\mathbf{I}_n$ is the $n \times n$ identity matrix. The dual form requires the inversion of an $n \times n$ matrix rather than a $p \times p$ matrix. Given that $n = 64$ examples and $p = 11{,}232$ features per example in our case, the dual form enables the algorithm to run considerably faster.

$\hat{\mathbf{W}}$ is a $p \times 3$ matrix where each column contains the weights for one of the three labels. These weights can be used to predict classes (e.g., Fan 1 or Fan 2) to which new examples belong by multiplying new examples $\mathbf{T}$ by $\hat{\mathbf{W}}$:

$$\hat{\mathbf{Y}} = \mathbf{T} \cdot \hat{\mathbf{W}} \tag{3}$$

$\hat{\mathbf{Y}}$ denotes the classification. In our case, the labels $\mathbf{Y}$ were either 1 or −1. Positive $\hat{\mathbf{Y}}$ were interpreted as a 1 classification (Fan 2, long, or target), and negative $\hat{\mathbf{Y}}$ as a −1 classification (Fan 1, short, or re-paired foil).

Training and Testing the Classifier

We trained and tested the classifier separately for each participant. This involved two steps. First, the best value of λ (the complexity parameter) was determined. Second, the classifier was trained and tested on separate sets of examples. For both steps, we used leave-one-out cross-validation (LOOCV). That is, we trained the classifier on 63 examples and used the resulting $\hat{\mathbf{W}}$ to classify the 64th example. We repeated this procedure for all examples, giving 64 accuracy measures.

To determine the optimal value for λ, we used LOOCV to minimize the mean squared prediction error (*MSE*) separately for each label (fan, length, and probe):

$$MSE = \frac{1}{n}\sum_{i=1}^{64}(y_i - \hat{y}_i)^2 \tag{4}$$

where y_i is the label for example i and $\hat{y}_i$ is the raw classification value for that example. We searched for a value of λ between 0 and 10,000 in the following manner. First, *MSE*s were calculated for 11 linearly spaced values of λ between 0 and 10,000. Next, we calculated *MSE*s for 11 linearly spaced points spanning the region that contained the value of λ that previously minimized *MSE*. We repeated this procedure until the improvement in *MSE* was smaller than .01. When the optimal value of λ was determined, we ran the classifier once more using LOOCV to calculate final prediction accuracy (percentage of correct classifications), which is reported in the Results section.

The classifier was first trained on all data between −200 and 1200 msec to assess how well it could perform given all data. To determine when different types of information processing occurred in the brain, we subsequently trained and tested the classifier using 50-msec windows. We included data from before stimulus presentation as a control; because the prestimulus interval does not contain information about condition, the classifier should perform at chance over this interval. To ensure that the stimulus- and response-locking procedure described above did not introduce latency confounds (i.e., conditions with shorter RTs were stretched more than conditions with longer RTs, affecting the alignment of classifier features and therefore possibly classifier performance), we also performed separate stimulus- and response-locked classifier analyses. For these analyses we excluded trials with an RT shorter than 700 msec and investigated intervals from −200 to 700 msec (stimulus-locked) and −700 to 200 msec (response-locked).

RESULTS

Behavioral Data

Training

The frequency with which target words were presented during each block of the cued recall task can be used to assess learning. The minimum possible frequency is one because each target word had to be presented at least once per block. The data were submitted to a repeated-measures ANOVA with Fan, Length, and Block as factors. Mean frequency decreased across blocks (4.4, 1.6, and 1.3 for Blocks 1–3, respectively), reflecting a main effect of Block, $F(2, 38) = 9.08$, $MSE = 26.60$, $p < .01$. The frequency was higher for Fan 2 pairs than for Fan 1 pairs (3.4 vs. 1.4, respectively), reflecting a main effect of Fan, $F(1, 19) = 15.48$, $MSE = 15.62$, $p < .01$, although this fan effect decreased across blocks, consistent with an interaction between Block and Fan, $F(2, 38) = 9.41$, $MSE = 10.83$, $p < .01$. No other effects were significant. The decreases in overall frequency and the fan effect across blocks both indicate progress in learning the target word pairs during the training phase.

Test

The test data were trimmed by excluding 2.3% of trials with RTs more than 3 *SD*s longer than the mean correct RT for a given condition and participant, leaving a mean of 98 observations per condition for each participant. The mean correct RTs and mean error rates appear in Table 1. RT was longer, and error rate was higher for Fan 2 probes compared with Fan 1 probes, reflecting main effects of fan on RT, $F(1, 19) = 67.30$, $MSE = 71131$, $p < .01$, and on error rate, $F(1, 19) = 26.57$, $MSE = 0.002$, $p < .01$, in a pair of repeated-measures ANOVAs with Probe, Fan, and Length as factors. RT was longer and the fan effect on RT was larger for re-paired foils than for targets, reflecting a main effect of probe, $F(1, 19) = 47.21$, $MSE = 17154$, $p < .01$, and an interaction between probe and fan, $F(1, 19) = 39.25$, $MSE = 4513$, $p < .01$. There were no other significant effects on either RT or error rate. These results indicate that re-paired foils were slightly more difficult compared with targets and that there were large effects of associative interference produced by the fan manipulation for both probe types.

Table 1. Mean Correct RTs (in msec) and Mean Error Rates (as %) in the Test Phase, with *SEM*s in Parentheses

		Fan 1		*Fan 2*	
Measure	*Probe*	*Short*	*Long*	*Short*	*Long*
RT	Target	1015 (38)	1026 (31)	1317 (67)	1282 (61)
	Re-paired Foil	1079 (45)	1113 (37)	1493 (82)	1524 (96)
Error rate	Target	3.7 (0.7)	2.5 (0.6)	7.3 (1.0)	6.1 (1.1)
	Re-paired Foil	3.1 (0.7)	3.1 (0.6)	5.8 (1.2)	6.1 (1.4)

ERP Data

Stimulus-locked Analysis

Analysis of the 32 electrodes and the 24 stimulus-locked time periods revealed significant effects of word length from 300 to 350 msec (see Figure 1), fan from 400 to 450 msec and from 700 to 900 msec (see Figure 2), and probe from 500 to 900 msec (see Figure 3). The word length effect was present at a single electrode, the early fan effect was present at 16 electrodes, the late fan effect was present at 5 electrodes, and the probe effect was present at 23 electrodes. Significant electrodes were contiguous in each temporally defined window and are denoted by points in Figures 1, 2, and 3. To examine these effects in detail, we performed separate 2 (Length) × 2 (Fan) × 2 (Probe) repeated-measures ANOVAs on data from ROIs during each of the four periods.

Word length produced a significant effect from 300 to 350 msec, $F(1, 19) = 35.67$, $MSE = 1.23$, $p < .001$. Long words produced more positive voltages than short words over the left-frontal scalp (see Figure 1). The Length effect was weaker over the central region and was absent over the parietal region. No other main effects or interactions reached significance.

Fan produced a significant effect from 400 to 450 msec, $F(1, 19) = 57.08$, $MSE = 1.07$, $p < .0001$. Fan 1 pairs yielded more negative voltages than Fan 2 pairs over the central and parietal scalp (see Figure 2, top). The main effect of Probe was also significant, $F(1, 19) = 13.48$, $MSE = 0.33$, $p < .01$, but the main effect of Length and the interactions were not.

Fan produced another significant effect from 700 to 900 msec, $F(1, 19) = 51.16$, $MSE = 1.66$, $p < .0001$, though in the reverse direction. Waveforms were more positive following Fan 1 pairs than following Fan 2 pairs. The late fan effect was maximal over the right-frontal and central scalp (see Figure 2, bottom). The main effect of Probe was also significant, $F(1, 19) = 27.28$, $MSE = 2.06$, $p < .0001$, but the main effect of Length and the interactions were not.

Probe produced a significant effect from 500 to 900 msec, $F(1, 19) = 55.45$, $MSE = 78.25$, $p < .0001$. Targets evoked more positive waveforms than re-paired foils over the right-parietal and central scalp (see Figure 3). The main effect of fan was also significant, $F(1, 19) = 6.76$, $MSE = 9.69$, $p < .05$, but the main effect of Length and the interactions were not.

Response-locked Analysis

We performed a response-locked analysis of the ERP data to investigate stimulus-related processing while controlling

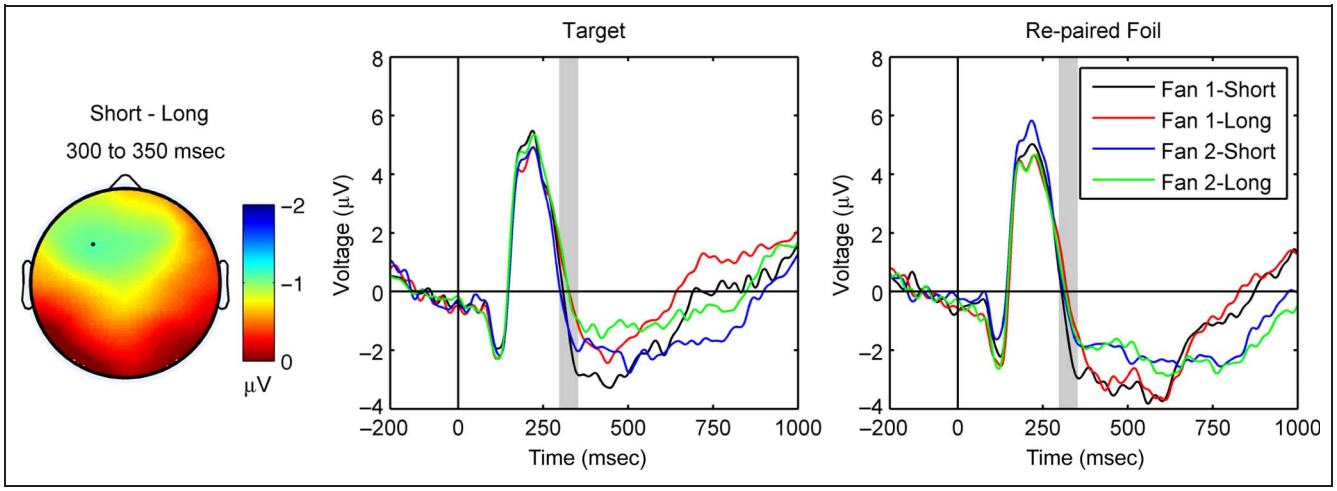


Figure 1. Topography and waveforms for stimulus-locked word length effect. Voltages differed between short and long words from 300 to 350 msec (site F3).

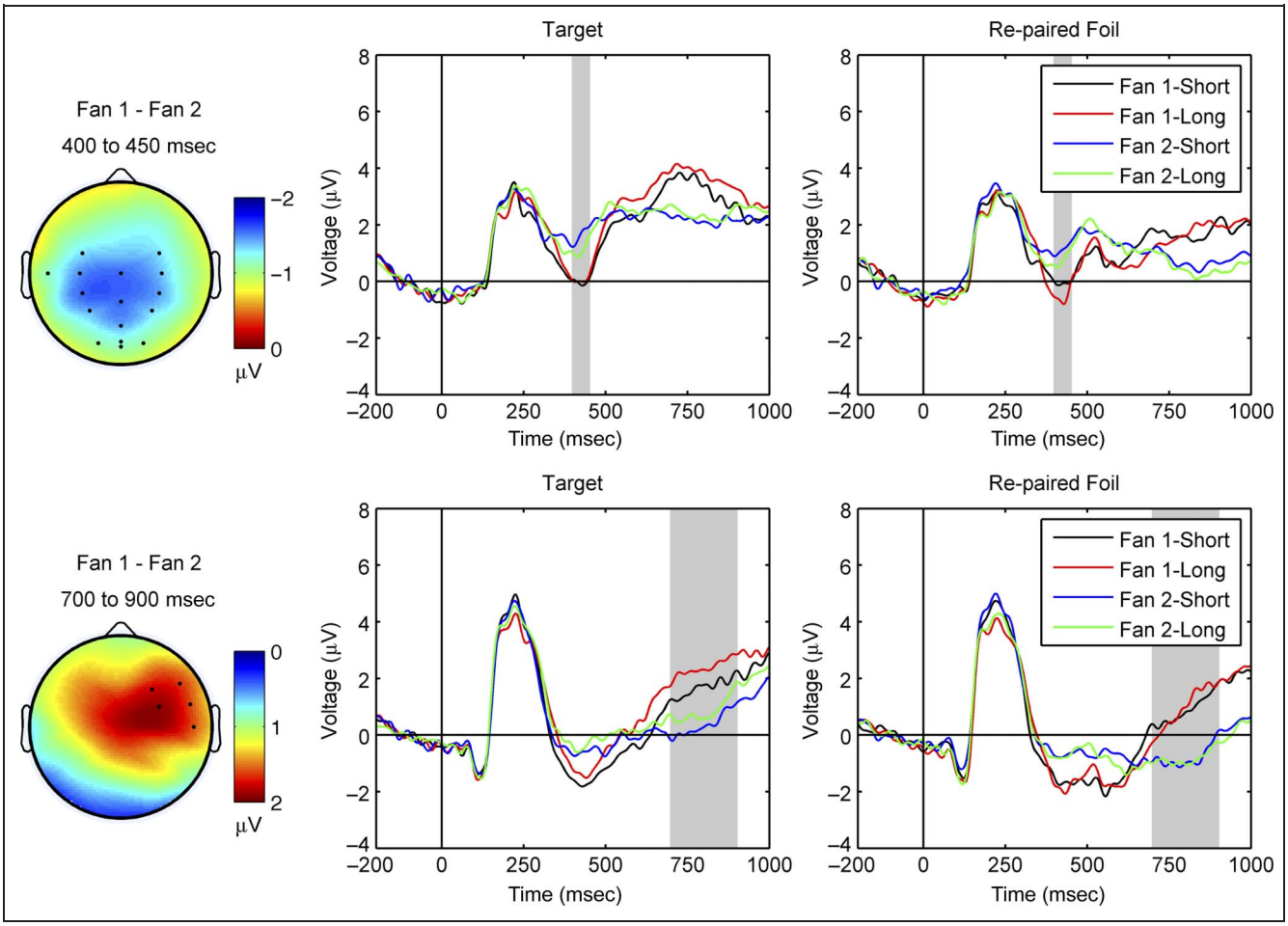


Figure 2. Topography and waveforms for stimulus-locked fan effects. Voltages differed between Fan 1 and Fan 2 pairs from 400 to 450 msec (top; sites FC3, FC4, T7, C3, Cz, C4, CP3, CPz, CP4, P3, Pz, P4, POz, O1, Oz, O2) and from 700 to 900 msec (bottom; sites F4, F8, FC4, FT8, T8). Note the reversal in the direction of the late fan effect compared with the early fan effect.

for response onset. Analysis of the 32 electrodes and the 24 response-locked time periods revealed no significant effects of length, but significant effects of fan from −150 to 0 msec (see Figure 4) and of probe from −350 to −50 msec (see Figure 5). The fan effect was present at 5 electrodes, and the probe effect was present at 20 electrodes. Significant electrodes were contiguous in each temporally defined window and are denoted by

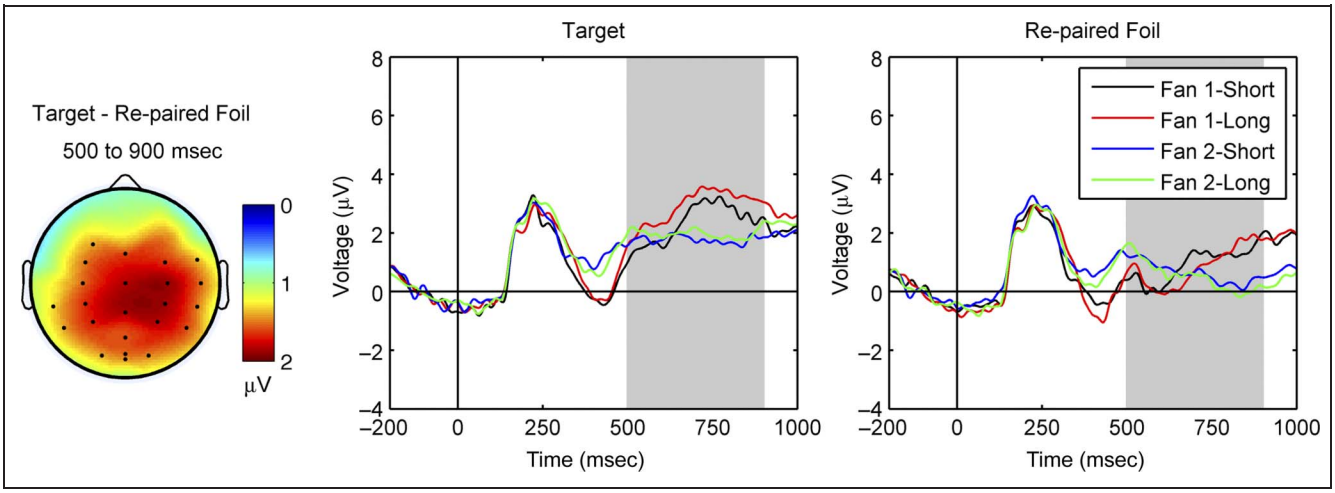


Figure 3. Topography and waveforms for stimulus-locked probe effect. Voltages differed between targets and re-paired foils from 500 to 900 msec (sites F3, FC3, FCz, FC4, FT8, C3, Cz, C4, T8, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, POz, O1, Oz, O2).

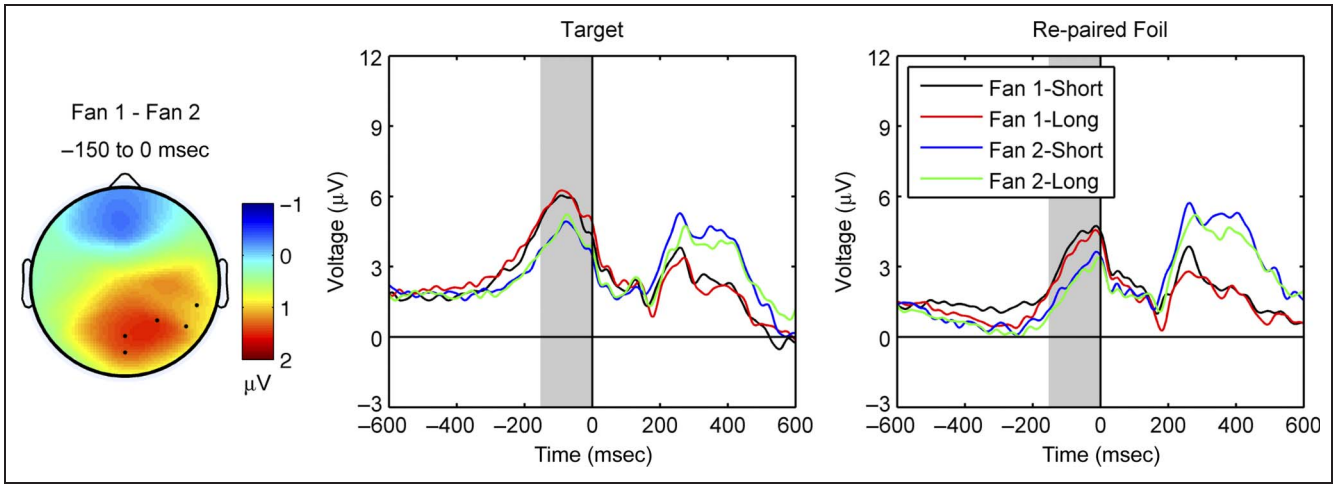


Figure 4. Topography and waveforms for response-locked fan effect. Voltages differed between Fan 1 and Fan 2 pairs from −150 to 0 msec (sites TP8, Pz, P4, P8, POz).

points in Figures 4 and 5. To examine these effects in detail, we performed separate 2 (Length) × 2 (Fan) × 2 (Probe) repeated-measures ANOVAs on the data from the two periods.

Fan produced a significant effect from −150 to 0 msec, $F(1, 19) = 23.82$, $MSE = 2.28$, $p < .001$. Waveforms were more positive after Fan 1 pairs than after Fan 2 pairs over the midparietal region (see Figure 4). The fan effect was weaker over the central region and was absent over the frontal region. The main effect of Probe was also significant, $F(1, 19) = 19.93$, $MSE = 5.99$, $p < .01$, but the main effect of Length and the interactions were not.

Probe produced a significant effect from −350 to −50 msec, $F(1, 19) = 48.38$, $MSE = 2.89$, $p < .0001$. Waveforms were more positive after targets than after re-paired foils over the midparietal and central scalp (see Figure 5). No other main effects of interactions were significant.

Classifier

Figure 6D shows classification accuracy when the classifier was trained and tested on all data, averaged over participants. The red bar indicates 37.7% accuracy for classifying all three conditions (fan, length, and probe) correctly on an example (chance is 12.5%). The blue, green, and orange bars indicate accuracy for classifying fan, length, and probe separately, which ranged from 67.3% for length to 79.2% for fan (chance is 50%). Accuracy was significantly greater than chance in all cases (ps < .001). Table 2 reports minimal and maximal *MSE*s over the range of complexity values that was applied during the classifier training. In addition, it shows the mean optimal complexity parameter.

Figure 6A shows classification accuracy when the classifier was trained and tested on 50-msec windows separately for fan, length, and probe. The horizontal bars at

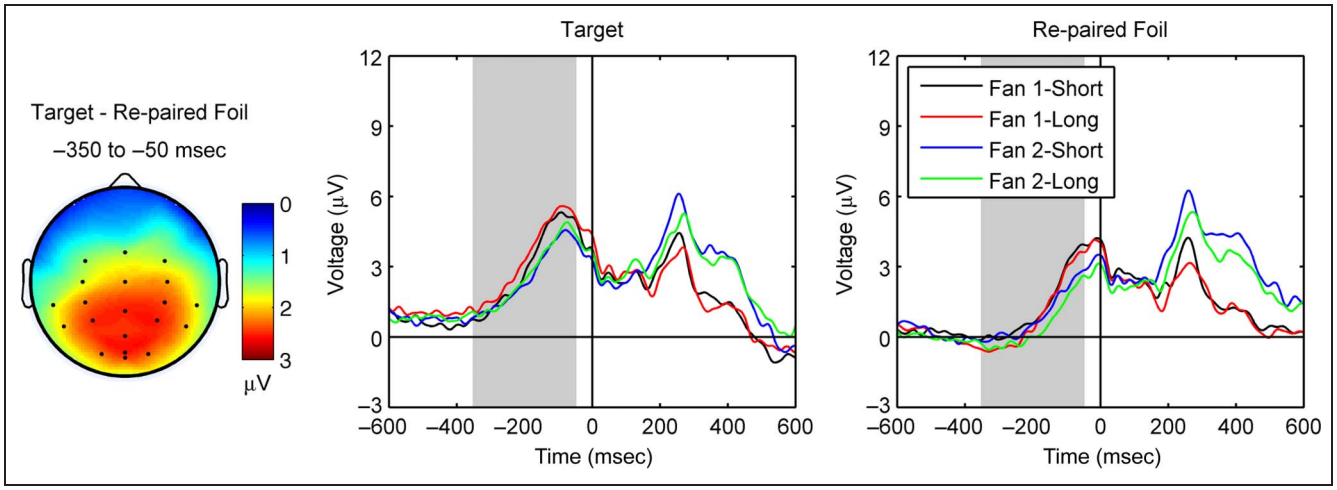


Figure 5. Topography and waveforms for response-locked probe effect. Voltages differed between targets and re-paired foils from −350 to −50 msec (sites FC3, FCz, FC4, C3, Cz, C4, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, POz, O1, Oz, O2).

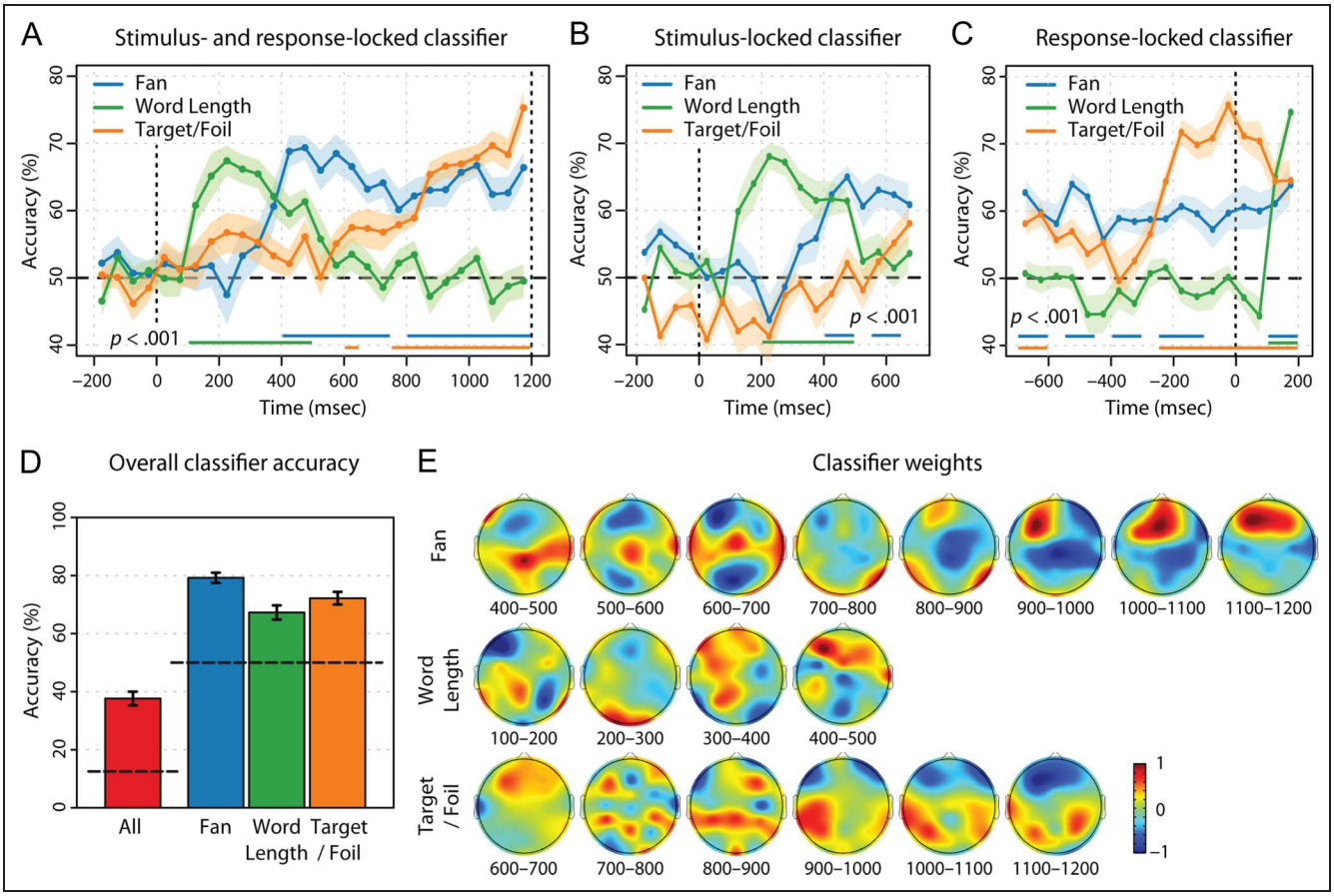


Figure 6. Classifier results. (A) Percentage of correctly classified examples when trained using 50-msec windows on stimulus- and response-locked data, where 0 denotes stimulus onset and 1200 the response (see the main text for details). The colored horizontal bars at the bottom indicate when accuracy significantly exceeded 50% ($p < .001$). The same information (B) for stimulus-locked data (C) for response-locked data (0 denotes the response). (D) Percentage of correctly classified examples when trained on both stimulus- and response-locked data, with the broken lines indicating chance. (E) Classifier weights corresponding to significant windows in A, averaged over 100-msec periods. Error bars in D and shaded areas in A–C denote *SEM*s.

the bottom of the figure indicate when classification accuracy significantly exceeded 50% (based on t tests for each window, a conservative threshold of $p < .001$ was used given the multiple comparisons). Before stimulus onset, the classifier performed at chance. The first variable classified above chance was length: The EEG signal contained information from 100 to 500 msec that enabled the classifier to distinguish between long and short words. The next variable classified above chance was fan: from 400 msec onward, the classifier distinguished between Fan 1 and Fan 2 pairs. The last variable classified above chance was probe: although the classifier distinguished between targets and re-paired foils beginning around 600 msec, classification accuracy continued to increase over time, reaching 75.3% just before the response.

Table 2. Minimal and Maximal Mean Squared Prediction Error and Mean Optimal Complexity Parameter

	Mean Squared Prediction Error		
	Minimum	*Maximum*	*Optimal* λ
Fan	0.60 (0.037)	0.70 (0.040)	2240 (549)
Word length	0.79 (0.032)	0.91 (0.042)	3975 (1234)
Probe	0.72 (0.039)	0.88 (0.063)	7092 (1272)

*SEM*s in parentheses.

Figure 6B and 6C show separate stimulus- and response-locked analyses, respectively. These analyses replicate the effects in Figure 6A.[3] The main difference was a later and more sudden onset of reliable probe classification in the response-locked analysis. Whereas the stimulus- and-response-locked classifier reliably distinguished between targets and re-paired foils from about 400 msec before the response, the response-locked classifier only distinguished between these conditions from about 300 msec before the response, which is in agreement with the ERP analysis.

Figure 6E shows the classifier weights ($\hat{\mathbf{W}}$) corresponding to significant windows of Figure 6A, averaged over 100-msec periods. These weights are the averages of the 64 classifiers trained for each participant in the LOOCV, averaged over participants. Note that Fan 2 pairs, long

words, and targets were coded as 1, resulting in positive weights when Fan 2 pairs, long words, and targets yielded larger voltages than Fan 1 pairs, short words, and re-paired foils, respectively. We scaled the weights to a maximum absolute value of one for display purposes. For length, there was an early occipital positivity (not evident in the conventional ERP analysis) from 200 to 300 msec, followed by a later frontal positivity from 300 to 500 msec. For fan, there was a midparietal positivity from 400 to 700 msec followed by a negativity from 800 to 1200 msec, mirroring the early and late fan effects seen in the ERP analysis. For probe type, there was a left-parietal positivity from 900 msec onward.

DISCUSSION

In this study we examined the processing stages involved in associative recognition. We conducted an EEG experiment that involved manipulations of word length, associative fan, and probe type, which were intended to tap the perceptual encoding, retrieval, and decision stages assumed by an ACT-R model of associative recognition. In the following sections, we discuss what the behavioral, ERP, and classifier results pertaining to each manipulation reveal about the processing that occurred, plus the implications for modeling associative recognition.

Word Length Effects and Perceptual Encoding

The manipulation of word length was expected to affect the perceptual encoding stage and to be manifest early in performance. There were no effects of length in the behavioral data, possibly because the manipulation's early locus led to any effects on RT being absorbed or washed out by later processing. The ERP analyses revealed that long words produced more positive voltages than short words over the left-frontal scalp from 300 to 350 msec (see Figure 1), consistent with a previous finding of broad positivity over the frontal region from 300 msec (Van Petten & Kutas, 1990). The classifier showed that information in the EEG signal distinguished between long and short words from 100 to 500 msec (see Figure 6). Given that perceptual encoding was a necessary first step in processing and word length is a perceptual feature, it makes sense that the classifier detected information early in the EEG signal for discrimination of word length. Collectively, these results indicate that word length had an effect on the early perceptual encoding stage but not on later stages.

These results—particularly the success of the classifier in isolating an early neural signature of word length over the occipital scalp—are consistent with the broader neuroscientific literature. EEG (Hauk & Pulvermüller, 2004), MEG (Sudre et al., 2012; Assadollahi & Pulvermuller, 2003), PET (Mechelli, Humphreys, Mayall, Olson, & Price, 2000), and fMRI studies (Indefrey et al., 1997) show that the visual cortex contributes to word processing and that responses from areas within the visual cortex are modulated by word length. Although the early occipital word length effect is consistent with what is known about the visual cortex, the later frontal word length effect is more difficult to interpret. Van Petten and Kutas (1990) suggested that the frontal effect is related to some linguistic difference between long and short words; however, they did not perform the requisite manipulations necessary to make any stronger claim. At present, it is unclear whether the frontal word length effect depends on orthographic, phonological, or semantic properties of long and short words.

Associative Fan Effects and Retrieval

The manipulation of associative fan was expected to affect the retrieval stage and be manifest in the middle of performance. The behavioral data revealed longer RTs and higher error rates for Fan 2 pairs compared with Fan 1 pairs (see Table 1), in accord with previous research and the ACT-R model (e.g., see Anderson, 2007; Anderson & Reder, 1999). The ERP data painted an interesting picture, showing distinct early and late fan effects. The stimulus-locked analysis revealed an early fan effect from 400 to 450 msec, during which time Fan 1 pairs produced more negative voltages than Fan 2 pairs over central and parietal sites (see Figure 2). The stimulus-locked analysis also revealed a late fan effect from 700 to 900 msec, during which time Fan 1 pairs produced more positive voltages than Fan 2 pairs over right frontal and central sites. The response-locked analysis showed that this late fan effect increased until the response (see Figure 4).

The early and late fan effects in the ERP data can be understood in terms of dual-process theories of recognition (Rugg & Curran, 2007). Familiarity and recollection, the two processes evoked in dual-process theories, are sensitive to fan (Nyhus & Curran, 2009). The fan effects in the current study might provide information about these processes. The early fan effect, which occurred soon after the word length effect, could reflect a familiarity signal from initial processing of the encoded probe words. The late fan effect, which occurred before the response, could reflect recollection of studied word pairs.

The nature of the late effect—Fan 2 pairs producing more negative voltages than Fan 1 pairs over the frontal scalp—is similar in some respects to the negative slow potential observed by Heil et al. (1996). Nyhus and Curran (2009) also reported an effect of fan on ERPs over the frontal scalp. Our effect was right-lateralized, whereas Nyhus and Curran (2009) and Heil et al. (1996) found that the effect of fan was maximal over the left hemisphere. Our experiment differs from Heil et al. (1996) in two important ways. First, we applied a high-pass filter (0.1 Hz) to our data whereas Heil et al. did not. Second, our task permitted participants to respond significantly faster. These differences may have limited our ability to detect the slow potential reported by Heil et al. (1996).

The difference in the lateralization of the fan effect might also reflect differences in task materials. Both Khader et al.

(2005, 2007) and Heil et al. (1996) found that the topographical distribution of the fan effect varied by stimulus modality. Although fMRI experiments with stimuli similar to those used in the current study have found the strongest effects of fan in the left pFC (Danker et al., 2008), Sohn et al. (2003) reported effects in the right pFC as well.

The classifier showed there was information in the EEG signal from 400 msec onward that distinguished between Fan 1 and Fan 2 pairs (see Figure 6). The fact that fan information became available near the end of the time period during which the classifier accurately discriminated between word length (100–500 msec) suggests that retrieval began once the probe words had been encoded, with no intervening stage. This inference is consistent with the results of a study by Anderson, Bothell, and Douglass (2004), who monitored eye movements during associative recognition and found evidence suggesting that retrieval did not begin until the probe word pair had been fully encoded. The persistence of good classification accuracy for fan over several hundred milliseconds suggests that each of the distinct fan effects in the ERP data provided diagnostic information to the classifier. Collectively, the behavioral, ERP, and classifier results support the identification of the associative fan manipulation with a retrieval stage in associative recognition.

Probe Effects, Retrieval, and Decision-making

The manipulation of probe type was expected to affect both the retrieval and the decision stages and be manifest late in performance, just before the response. The behavioral data revealed longer RTs for re-paired foils compared with targets (see Table 1), which is consistent with the idea that it may be more difficult to make a decision when there is a match at the item level but not at the associative level (i.e., the probe words matched studied words but were from different pairs).

The stimulus-locked analysis of the ERP data revealed that targets produced more positive voltages than re-paired foils over parietal and central sites from 500 to 900 msec (see Figure 3), similar in some respects to the parietal old/new effect observed in previous research (e.g., Curran, 2000; Düzel et al., 1997). The response-locked analysis revealed that the effect was present just before the response (see Figure 5). The classifier showed that information in the EEG signal starting around 300 msec before the response distinguished clearly between targets and re-paired foils (see Figure 6; note that there seems to be an earlier effect around 600 msec before the response) and became increasingly diagnostic leading up to the response. The later onset of good classification accuracy for the probe conditions relative to the early fan effect indicates that targets and re-paired foils could only be differentiated once some retrieval had occurred.

As hypothesized, the probe manipulation resulted in a parietal old/new effect, which overlapped with the effect of fan during the retrieval stage. This effect was bilateral, although the parietal old/new effect is typically left lateralized (e.g., Rugg & Curran, 2007; however, the effect is often present over the right hemisphere as well, e.g., Woodruff, Hayama, & Rugg, 2006; Curran, 2000, 2004). One difference between previous studies of recognition memory and the current study is that old/new status differed between targets and re-paired foils at the associative level of word pairs rather than at the item level of individual words. In a study more similar to our own, Speer and Curran (2007) tested participants' ability to distinguish between studied and rearranged item pairs and also found strongest effects over the right hemisphere. Thus, associative differences in old/new status may result in stronger effects in the right hemisphere than in the left hemisphere.

In the current experiment, both targets and re-paired foils were presented 13 times during the test phase. Although participants initially had to perform associative recognition by remembering whether words had been studied together, they might have eventually switched to a strategy in which they determined when during the experiment (training vs. test phase) the words had been presented together. Although there was still a clear difference in RT between targets and re-paired foils in the last block of the experiment (1117 vs. 1277 msec), $F(1, 19) = 29.81, p < .001$, this might be because of differences in memory strength. To investigate whether the parietal old/new effect changed over the course of the experiment, we analyzed the effect of probe type from 500 to 900 msec during the first and second halves of the test phase. Targets consistently evoked a more positive response than re-paired foils over the right-parietal scalp and the interaction between probe and test half did not approach significance, $F(1, 19) = 0.97, p > .1$.

Implications for Modeling Associative Recognition

The behavioral, ERP, and classifier results generally support the nature and ordering of information processing stages assumed by the ACT-R model for associative recognition. The manipulations of word length, associative fan, and probe type affected the hypothesized stages of perceptual encoding, retrieval, and decision-making in the expected order and at sensible times, as reflected by modulations in the ERP waveforms and changes in classifier accuracy. However, the reversal of the fan effect in the ERP data during the time course of task performance implicates a more complex retrieval stage than originally assumed by the ACT-R model, which would predict a single fan effect in the latter half of the experiment.

We hypothesize that the early fan effect reflects access to item information in memory (i.e., whether the individual probe words were studied), whereas the late fan effect reflects access to associative information in memory (i.e., whether the probe word pair was studied). From the perspective of dual-process theories of recognition,

item information would become available because of familiarity and associative information would become available because of recollection (Diana et al., 2006; Yonelinas, 2002). The ACT-R model does not include a familiarity process. Item information is accessed in a manner analogous to how associative information is retrieved by using individual probe words to retrieve studied words from memory (Anderson, Bothell, Lebiere, & Matessa, 1998).

The retrieval of both item and associative information in the ACT-R model could be reconciled with the reversed fan effects in the ERP data if those effects reflect the strength of the memories being retrieved. In the case of item retrieval, Fan 2 words could have been represented more strongly in memory compared with Fan 1 words because they occurred more often among studied pairs, reflecting the nature of the fan manipulation. It also took participants longer to learn Fan 2 pairs during training, which meant they were exposed to Fan 2 words more often than Fan 1 words in that phase. Items presented more often during study are recognized better than items presented less often (e.g., Stretch & Wixted, 1998; Ratcliff, Clark, & Shiffrin, 1990). As a result, item retrieval for Fan 2 words might have been facilitated relative to that for Fan 1 words, resulting in a more positive potential for Fan 2 words. This is consistent with a study by Finnigan, Humphreys, Dennis, and Geffen (2002), who reported more positive N400 effects for strongly encoded items than for weakly encoded items. In the case of associative retrieval, Fan 2 pairs would be difficult to access than Fan 1 pairs because their constituent words were shared across multiple pairs, yielding a form of associative interference. As a result, associative retrieval for Fan 2 pairs would be impaired relative to that for Fan 1 pairs. Combining these ideas, facilitated item retrieval for Fan 2 words and impaired associative retrieval for Fan 2 pairs could have resulted in the reversed fan effects we observed in the ERP data.

The ERP and classifier results also suggest that the retrieval and decision stages coincide to some extent. The late fan effect persisted until the response (see Figures 4 and 6) and the response-locked analysis revealed similar scalp topographies for the late fan effect and the target-foil effect (see Figures 4 and 5). One interpretation of these results is that associative retrieval represents part of the decision process, with the outcome of retrieval automatically providing evidence for and against different decisions. Alternatively, these late parietal effects might indicate response confidence (e.g., Woodruff et al., 2006). Targets produced more positive voltages than re-paired foils, and Fan 1 pairs produced more positive voltages than Fan 2 pairs. These effects resemble P300 effects (the parietal old/new effect encompasses the P300; e.g., Curran, DeBuse, Woroch, & Hirshman, 2006), which are known to correlate with response confidence (e.g., Nieuwenhuis, Aston-Jones, & Cohen, 2005; Sutton, Ruchkin, Munson, Kietzman, & Hammer, 1982; Wilkinson & Seales, 1978).

Besides the ACT-R model, these results have implications for other models of associative recognition, including those that do not make the same distinctions among processing stages. For example, evidence accumulation models typically involve a stage that integrates the retrieval and decision processes (e.g., Ratcliff & Smith, 2004; Ratcliff, 1978). More specifically, evidence for responses gradually accumulates based on information retrieved from memory, and a decision occurs when the accumulated evidence reaches a criterion. The decision is not a separate process but rather the endpoint of retrieval-based evidence accumulation, which is consistent with the aforementioned similarities between the late fan effect and the probe effect.

Although it is possible for evidence accumulation models to produce the behavioral fan effects we observed, it is not obvious whether they could be reconciled with the reversed fan effects in the ERP data. One could argue that the reversed fan effects reflect changes in the type of evidence being accumulated, with the early fan effect reflecting item information and the late fan effect reflecting associative information, as discussed earlier. Indeed, there are models that allow for such changes in evidence accumulation during the time course of processing (e.g., Brockdorff & Lamberts, 2000; Ratcliff, 1980). However, given that both types of information would contribute to the same evidence accumulation process, one might expect the early and late fan effects to have similar scalp topographies. In contrast, we found that the early fan effect was manifest over left-central and left-parietal sites, whereas the late fan effect was manifest over right-frontal and right-central sites (see Figure 2).

Despite the challenges from the ERP data, evidence accumulation models do receive support from the classifier results (see Figure 6). A fan effect would be expected for the duration of the evidence accumulation process, consistent with the classifier's accuracy at distinguishing between Fan 1 and Fan 2 pairs from 400 msec onward. A probe effect would be expected to emerge gradually as the accumulated evidence points increasingly to either a target or a re-paired foil, consistent with the classifier's growing accuracy at distinguishing between targets and re-paired foils from 600 msec onward. However, note that the classification patterns are also consistent with the ACT-R model. Its discrete processing stages can produce graded changes in accuracy over time if the finishing times of the processes are variable (e.g., Schneider & Anderson, 2012). Thus, the classifier results do not uniquely support either model.

Conclusion

The preceding discussion highlights the ways in which the combination of behavioral, ERP, and classifier results can inform the modeling of associative recognition. Although the results provide general support for the stages of processing assumed by the ACT-R model, there was no prior

guarantee that they would. For example, from the perspective of a simple evidence accumulation model, one might have expected fan and probe effects in the ERP data to completely overlap temporally and topographically based on their attribution to a common stage. There was evidence of overlap between the late fan effect and the probe effect, but the early fan effect had different characteristics, suggesting a complex relationship between retrieval and decision-making processes. The classification methodology used in this study showed how machine-learning techniques can provide additional leverage in data interpretation. For example, although the standard ERP analysis did not reveal an early occipital effect of word length, the multivariate pattern used by the classifier enabled discrimination between word lengths as early as 100 msec. In addition, the classifier's discrimination of the fan and probe conditions provided insight regarding the retrieval and decision stages of processing that could not be inferred from the behavioral data alone. Thus, this study represents an example of how the detailed temporal data from EEG, supplemented by modern classification techniques, can be used to evaluate basic components of computational models of cognition.

Reprint requests should be sent to Jelmer P. Borst, Department of Psychology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, or via e-mail: jelmer@cmu.edu.

Notes

1. New foils were included for the initial purpose of contrasting item recognition with associative recognition. However, considering the many ways in which new foils differed from the other probe types (e.g., they were not present during the training phase, they were not repeated during the test phase, they could be rejected as nonstudied without retrieving associative information, and their associative fan could not be manipulated), we subsequently chose to focus exclusively on associative recognition and restrict our analyses to targets and re-paired foils.
2. Genovese et al. (2002) recommended setting the FDR between 0.01 and 0.05. Because of the exploratory nature of our study, we favored a conservative FDR (0.01). The results remained essentially unchanged as we varied the FDR over two orders of magnitude (from 0.001 to 0.1).
3. Trials with RTs shorter than 700 msec were excluded from these analyses, resulting in less power. Consequently, the length effect from 100 to 200 msec did not reach significance in the stimulus-locked classifier.

REFERENCES

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6,* 451–474.
Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.
Anderson, J. R., Betts, S., Ferris, J. L., & Fincham, J. M. (2011). Tracking children's mental states while solving algebra equations. *Human Brain Mapping, 33,* 2650–2665.
Anderson, J. R., Bothell, D., & Douglass, S. (2004). Eye movements do not reflect retrieval processes: Limits of the eye-mind hypothesis. *Psychological Science, 15,* 225–231.
Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38,* 341–380.
Anderson, J. R., Carter, C. S., Fincham, J. M., Qin, Y., Ravizza, S. M., & Rosenberg-Lee, M. (2008). Using fMRI to test models of complex cognition. *Cognitive Science, 32,* 1323–1348.
Anderson, J. R., Fincham, J. M., Schneider, D. W., & Yang, J. (2012). Using brain imaging to track problem solving in a complex state space. *Neuroimage, 60,* 633–643.
Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General, 128,* 186–197.
Assadollahi, R., & Pulvermuller, F. (2003). Early influences of word length and frequency: A group study using MEG. *NeuroReport, 14,* 1183–1187.
Brockdorff, N., & Lamberts, K. (2000). A feature-sampling account of the time course of old-new recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 77–102.
Chan, A. M., Halgren, E., Marinkovic, K., & Cash, S. S. (2011). Decoding word and category-specific spatiotemporal representations from MEG and EEG. *Neuroimage, 54,* 3028–3039.
Coles, M. G. H. (1988). Modern mind-brain reading: Psychophysiology, physiology, and cognition. *Psychophysiology, 26,* 251–269.
Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A,* 497–505.
Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition, 28,* 923–938.
Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia, 42,* 1088–1106.
Curran, T., DeBuse, C., Woroch, B., & Hirshman, E. (2006). Combined pharmacological and electrophysiological dissociation of familiarity and recollection. *Journal of Neuroscience, 26,* 1979–1985.
Danker, J. F., Gunn, P., & Anderson, J. R. (2008). A rational account of memory predicts left prefrontal activation during controlled retrieval. *Cerebral Cortex, 18,* 2674–2685.
Das, K., Giesbrecht, B., & Eckstein, M. P. (2010). Predicting variations of perceptual performance across individuals from neural activity using pattern classifiers. *Neuroimage, 51,* 1425–1437.
Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134,* 9–21.
Diana, R. A., Reder, L. M., Arndt, J., & Park, H. (2006). Models of recognition: A review of arguments in favor of a dual process account. *Psychonomic Bulletin & Review, 13,* 1–21.
Dien, J. (2009). The neurocognitive basis of reading single words as seen through early latency ERPs: A model of converging pathways. *Biological Psychology, 80,* 10–22.
Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30,* 412–431. (Translation of original work published in 1868.)
Düzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H. J., & Tulving, E. (1997). Event-related brain potential correlates of two stages of conscious awareness in memory. *Proceedings of the National Academy of Sciences, U.S.A., 94,* 5973–5978.

Finnigan, S., Humphreys, M. S., Dennis, S., & Geffen, G. (2002). ERP "old/new" effects: Memory strength and decisional factor(s). *Neuropsychologia, 40,* 2288–2304.
Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage, 15,* 870–878.
Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning. Data mining, inference, and prediction*. New York: Springer-Verlag.
Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage, 30,* 1383–1400.
Hauk, O., & Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology, 115,* 1090–1103.
Hauk, O., Pulvermüller, F., Ford, M., Marslen-Wilson, W. D., & Davis, M. H. (2009). Can I have a quick word? Early electrophysiological manifestations of psycholinguist processes revealed by event-related regression analysis of the EEG. *Biological Psychology, 80,* 64–74.
Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience, 7,* 523–534.
Heil, M., Rösler, F., & Hennighausen, E. (1996). Topographically distinct cortical activation in episodic long-term memory: The retrieval of spatial versus verbal information. *Memory & Cognition, 24,* 777–795.
Heil, M., Rösler, F., & Hennighausen, E. (1997). Topography of brain electrical activity dissociates the retrieval of spatial versus verbal information from episodic long-term memory in humans. *Neuroscience Letters, 222,* 45–48.
Indefrey, P., Kleinschmidt, A., Merboldt, K. D., Kruger, G., Brown, C., Hagoort, P., et al. (1997). Equivalent responses to lexical and nonlexical visual stimuli in occipital cortex: A functional magnetic resonance imaging study. *Neuroimage, 5,* 78–81.
Khader, P., Heil, M., & Rösler, F. (2005). Material-specific long-term memory representations of faces and spatial positions: Evidence from slow event-related brain potentials. *Neuropsychologia, 43,* 2109–2124.
Khader, P., Knoth, K., Burke, M., Ranganath, C., Bien, S., & Rösler, F. (2007). Topography and dynamics of associative long-term memory retrieval in humans. *Journal of Cognitive Neuroscience, 19,* 493–512.
Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
Mechelli, A., Humphreys, G. W., Mayall, K., Olson, A., & Price, C. J. (2000). Differential effects of word length and visual contrast in the fusiform and lingual gyri during reading. *Proceedings of the Royal Society of London, Series B, 267,* 1909–1913.
Mitchell, T. M., Shinkareva, S., Carlson, A., Chang, K., Malave, V., Mason, R., et al. (2008). Predicting human brain activity associated with the meanings of nouns. *Science, 320,* 1191–1195.
Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin, 131,* 510–532.
Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences, 10,* 424–430.
Nyhus, E., & Curran, T. (2009). Semantic and perceptual effects on recognition memory: Evidence from ERP. *Brain Research, 1283,* 102–114.
Parra, L., Alvino, C., Tang, A., Pearlmutter, B., Yeung, N., Osman, A., et al. (2002). Linear spatial integration for single-trial detection in encephalography. *Neuroimage, 17,* 223–230.
Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage, 45,* S199–S209.
Peters, B. O., Pfurtscheller, G., & Flyvbjerg, H. (1998). Mining multi-channel EEG for its information content: An ANN-based method for a brain-computer interface. *Neural Networks, 11,* 1429–1433.
Philiastides, M. G., & Sajda, P. (2006). Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cerebral Cortex, 16,* 509–518.
Pirolli, P. L., & Anderson, J. R. (1985). The role of practice in fact retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 136–153.
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.
Ratcliff, R. (1980). A note on modeling accumulation of information when the rate of accumulation changes over time. *Journal of Mathematical Psychology, 21,* 178–184.
Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 163–178.
Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111,* 333–367.
Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences, 11,* 251–257.
Schneider, D. W., & Anderson, J. R. (2012). Modeling fan effects on the time course of associative recognition. *Cognitive Psychology, 64,* 127–160.
Smulders, F. T. Y., & Miller, J. O. (2013). The lateralized readiness potential. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 209–229). New York: Oxford University Press.
Sohn, M. H., Goode, A., Stenger, V. A., Carter, C. S., & Anderson, J. R. (2003). Competition and representation during memory retrieval: Roles of prefrontal cortex and the posterior parietal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 100,* 7412–7417.
Speer, N. K., & Curran, T. (2007). ERP correlates of familiarity and recollection processes in visual associative recognition. *Brain Research, 1174,* 97–109.
Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30,* 276–315.
Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1379–1396.
Sudre, G., Pomerleau, D., Palatucci, M., Wehbe, L., Fyshe, A., Salmelin, R., et al. (2012). Tracking neural coding of perceptual and semantic features of concrete nouns. *Neuroimage, 62,* 451–463.
Sutton, S., Ruchkin, D. S., Munson, R., Kietzman, M. L., & Hammer, M. (1982). Event-related potentials in a two-interval forced-choice detection task. *Psychonomic Bulletin & Review, 32,* 360–374.
Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition, 18,* 380–393.

Vilberg, K. L., Moosavi, R. F., & Rugg, M. D. (2006). The relationship between electrophysiological correlates of recollection and amount of information retrieved. *Brain Research, 1122,* 161–170.
Vilberg, K. L., & Rugg, M. D. (2009). Functional significance of retrieval-related activity in lateral parietal cortex: Evidence from fMRI and ERPs. *Human Brain Mapping, 30,* 1490–1501.
Wilding, E. L. (2000). In what way does the parietal ERP old/new effect index recollection? *International Journal of Psychophysiology, 35,* 81–87.
Wilkinson, R. T., & Seales, D. M. (1978). EEG event-related potentials and signal detection. *Biological Psychology, 7,* 13–28.
Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research, 1100,* 125–135.
Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441–517.
Zipf, G. K. (1935). *The psycho-biology of language*. Boston: Houghton Mifflin.