Event Congruency Enhances Episodic Memory Encoding through Semantic Elaboration and Relational Binding

Bernhard P. Staresina[1], James C. Gray[1] and Lila Davachi[1,2]

[1]Department of Psychology and [2]Center for Neural Science, New York University, New York, NY 10003, USA

Behavioral research consistently shows that congruous events, that is, events whose constituent elements match along some specific dimension, are better remembered than incongruous events. Although it has been speculated that this "congruency subsequent memory effect" (cSME) results from enhanced semantic elaboration, empirical evidence for this account is lacking. Here, we report a set of behavioral and neuroimaging experiments demonstrating that congruous events engage regions along the left inferior frontal gyrus (LIFG)—consistently related to semantic elaboration—to a significantly greater degree than incongruous events, providing evidence in favor of this hypothesis. Critically, we additionally report 3 novel findings in relation to event congruency: First, congruous events yield superior memory not only for a given study item but also for associated source details. Second, the cSME is evident not only for events that matched a semantic context but also for those that matched a subjective aesthetic schema. Finally, functional magnetic resonance imaging brain/behavior correlation analysis reveals a strong link between 1) across-subject variation in the magnitude of the cSME and 2) differential right hippocampal activation, suggesting that episodic memory for congruous events is effectively bolstered by the extent to which semantic associations are generated and relationally integrated via LIFG-hippocampal–encoding mechanisms.

Keywords: episodic memory, fMRI, hippocampus, prefrontal cortex, semantic memory

Introduction

A fundamental approach to understanding memory has been to ask how different encoding variables affect the formation of a new memory trace. One influential framework, the levels of processing (LOP) framework (Craik and Lockhart 1972), asserts that the strength of a memory trace is a function of the "depth" of encoding, "[…] where depth refers to greater degrees of semantic involvement" (Craik and Tulving 1975). Behavioral research has provided compelling evidence in support of the LOP framework (Lockhart and Craik 1990), and more recently, neuroimaging studies have shed some light on the neurophysiological substrates of different LOP (Demb et al. 1995; Wagner et al. 1998; Otten et al. 2001). Specifically, Wagner et al. (1998) revealed that regions along the left inferior frontal gyrus (LIFG) were more active during "deep" semantic encoding (deciding if a given word is abstract or concrete) than during "shallow" nonsemantic encoding (deciding if a given word is presented in upper- or lowercase letters). Furthermore, in the same study, it was found that the magnitude of encoding activation in these brain regions correlated with successful memory formation, that is, predicted whether a given study item would later be remembered or forgotten (subsequent memory effect or SME). These results provided strong converging evidence for a relationship between semantic encoding and successful memory formation.

In addition to the LOP, however, "event congruency" has been shown to strongly influence memory performance. Specifically, congruous encoding events eliciting "yes" answers (e.g., "is CAR a manmade object?") have been found to result in better subsequent recognition and recall memory than incongruous events eliciting "no" answers (e.g., "is CAR a natural object?") (Craik and Tulving 1975). A very similar effect was reported by Schulman (1974), who also presented subjects with congruous (e.g., "is a corkscrew an opener?") and incongruous (e.g., "is a dungeon a scholar?") events and found that later memory performance was superior for the elements of congruous events (i.e., memory was better for "corkscrew" and "opener" compared with "dungeon" and "scholar"). We henceforth refer to enhanced subsequent memory for congruous relative to incongruous encoding events as the "congruency subsequent memory effect" or cSME.

Despite the intriguing impact of event congruency on subsequent memory performance (see also Hall and Geis 1980), the underlying mechanisms of this effect have remained underspecified. Both Schulman (1974) and Craik and Tulving (1975) speculated that congruous events foster additional "semantic elaboration." More precisely, they argued that the elements of congruous events, unlike incongruous ones, form an integrated unit that has a preexisting semantic relationship. This semantic relationship is thought to prompt additional semantic-associative "spread" during encoding, which in turn renders the memory trace more accessible for subsequent memory tests. However, without another experimental marker of semantic elaboration, this rationale is somewhat circular: 1) semantic elaboration has been shown to benefit later memory and 2) event congruency enhances subsequent memory performance, thus 3) this memory advantage must be driven by enhanced semantic elaboration during congruous encoding events. But behavioral evidence for enhanced semantic elaboration has been lacking, and potentially related variables like reaction time data do not show a difference between congruous- ("yes") and incongruous- ("no") encoding events (Craik and Tulving 1975). Moreover, Moscovitch and Craik (1976) raised the argument that the mnemonic advantage of congruous events may emerge at the stage of retrieval rather than being due to differential encoding operations. That is, they suggest that a retrieval cue is more effective when being part of a congruous cue–target combination because the target

is—by definition of "congruency"—more readily available for those cues. Accordingly, event congruency would not necessarily affect the way a target is initially encoded but exert its beneficial mnemonic effect by providing a more effective cue–target relationship during memory retrieval.

In short, not only has the account that event congruency enhances subsequent memory performance through semantic elaboration been challenged on theoretical grounds but empirical evidence in support of this account has not been provided so far. In this case, functional neuroimaging data can be one means through which empirical support is provided. If the cSME is driven by semantic elaboration, congruous events should differentially engage regions that are 1) typically involved in semantic elaboration and 2) related to successful memory formation.

In this paper, we report a set of experiments designed to further investigate the cSME and elucidate its underlying mechanisms both with behavioral and neuroimaging measures. First, we used functional magnetic resonance imaging (fMRI) to assess the neural mechanisms underlying the cSME. On a mechanistic level, regions underlying the cSME should 1) be more engaged during congruous relative to incongruous events and 2) globally support successful memory formation. Second, on a qualitative level, if event congruency enhances subsequent memory through semantic elaboration, the resulting regions should overlap with those typically involved in semantic processing, such as the LIFG (Wagner et al. 1998). However, how semantic elaboration per se benefits subsequent memory performance remains unclear. That is, although congruency-induced semantic elaboration may serve to activate additional associations as proposed by Schulman (1974) and Craik and Tulving (1975), we would argue that episodic memory for the encoding event should ultimately depend on the extent to which those additional associations are integrated or bound to form a robust mnemonic representation, perhaps reflected in enhanced recruitment of hippocampal relational encoding operations (Cohen and Eichenbaum 1993; Davachi and Wagner 2002; Davachi et al. 2003; Ranganath et al. 2004).

Second, we employed 2 different encoding tasks that would allow us to query behaviorally whether the cSME is restricted to events where congruency denotes a purely semantic match between the constituent elements: While subjects were asked to rate the perceived semantic match of specific word/color combinations in a "plausibility" task (e.g., is a "red elephant" plausible?), they were asked to rate the perceived aesthetic match of specific word/color combinations in a "valence" task (e.g., is a "blue shirt" appealing?; Fig. 1A). Event congruency was thus operationalized as trials where word and color were given "plausible" judgments in the plausibility task or "appealing" judgments in the valence task, respectively. Next, we assessed whether the cSME is restricted to superior subsequent memory for the concrete elements of an encoding event or, rather, extends to include enhanced memory for associated source details such as the surrounding encoding context (the specific encoding task during which a word was encountered). That is, if semantic elaboration is indeed enhanced during congruous events, this begs the question of whether the cSME will, in addition to showing enhanced memory for the specific components of an event (corkscrew and opener in the above example), also show enhanced memory for contextual source details associated with the event (Johnson et al. 1993).
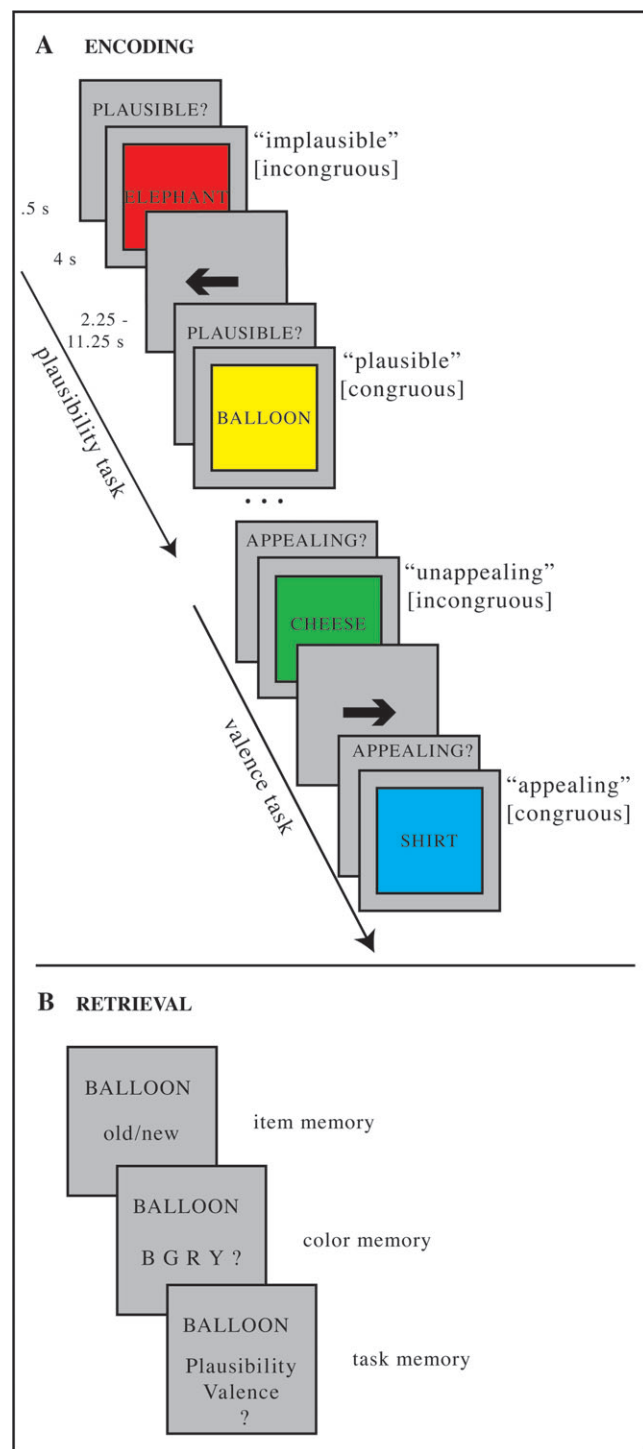

Figure 1. Design for Experiments 1 and 2. (A) During encoding (scanned with fMRI in Experiment 2), subjects were presented with word/color combinations in blocks of a plausibility task alternating with blocks of a valence task. For a given trial, subjects were asked to vividly imagine the referent of the noun in the given color and to either rate the plausibility of the given word/color combination to appear in real life/nature (plausible or implausible in the plausibility task) or the subjective aesthetic appeal of the word/color combination (appealing or unappealing in the valence task). Example trials show possible answers a subject may give for the given word/color combinations and (in brackets) the resulting classification of that event as congruous or incongruous. (B) The following 3-step surprise recognition memory test assessed item memory by asking subjects to make old/new judgments on previously seen or unseen words. If the response was old, relational/source memory was assessed by testing memory for 1) the associated color with which an item was presented and 2) the associated task context in which an item was encountered.

Materials and Methods

Subjects and Material

A total of 39 right-handed native English speakers (16 males) participated across 3 experiments. Mean age across subjects was 20 years with a range from 18 to 27. All subjects had normal or corrected-to-normal vision. Written informed consent was obtained in a manner approved by the institutional review board at New York University and subjects were paid for their participation. In all, 8 subjects participated in Experiment 1 (an additional 3 subjects were excluded from all analyses due to early termination of the experiment), 22 subjects participated in Experiment 2 (one additional subject was excluded from all analyses for providing only 1 trial of successful task encoding), and the remaining 9 subjects participated in Experiment 3. For Experiments 1 and 2, the item material consisted of 700 English nouns referring to concrete objects. The item pool was counterbalanced so that across subjects, every word was presented during both encoding and retrieval and was presented with every color in both encoding tasks. For Experiment 3, a subset of 432 items from this item pool was used.

Behavioral Procedures

Experiments 1 and 2

During encoding (Fig. 1A), subjects were presented with 420 word/color combinations that were divided into 5 blocks of a plausibility task alternating with 5 blocks of a valence task, each block containing 42 trials. The sequence of task blocks was A-B-B-A-A-B-B-A-A-B, with the assignment of the particular task to A and B being counterbalanced across subjects. In Experiment 1, encoding trials were separated by a 2-s-long fixation cross. In Experiment 2 (fMRI study), encoding trials were intermixed with baseline trials of an active, sensorimotor "arrows" task (Stark and Squire 2001). Arrows that randomly pointed to the left or to the right for 1 s were repeatedly presented for the length of a baseline trial, and subjects had to press the left middle finger key if the arrow pointed to the left and the left index finger key if it pointed to the right. The order of word/color trials and baseline trials was determined by using a sequencing program designed to maximize the efficiency of the event-related design (Dale 1999). Conditions were jittered using variable duration (2.25–11.25 s) baseline trials.

For a given trial (4.5 s total), subjects were presented with a 0.5-s-long task cue ("plausible?" in the plausibility task and "appealing?" in the valence task) followed by the 4-s-long presentation of a noun superimposed on a color square (blue, green, red, or yellow). For each trial, subjects were instructed to vividly imagine the referent of the noun in the presented color and to make a specific judgment based on the current encoding task. In the plausibility task, subjects were asked to indicate whether it was plausible to encounter the imagined object/color combination in real life/nature or not, with the answer options being "plausible" or "implausible." In the valence task, subjects were asked to indicate whether they thought the imagined object/color combination was aesthetically "appealing" or not, with the answer options being appealing or "unappealing." Note that for both tasks, there was not necessarily an objectively correct or incorrect answer for every word/color combination. That is, the perceived plausibility of the combination STRAWBERRY/green in the plausibility task would depend on whether a given subject may come up with an image of an unripe strawberry (in which case he or she would give a "plausible" response) or whether he or she thinks strawberries ought to be red, evoking an "implausible" response for that particular trial. Likewise, in the valence task, whereas one subject may feel that the combination SHIRT/blue is appealing, another subject might feel it is unappealing. Thus, both encoding tasks required subjective ratings of the given word/color combinations, but the criterion varied between plausible in real life/nature in the plausibility task and subjectively appealing in the valence task. Event congruency was thus defined by subjects themselves by rating trials as plausible in the plausibility task and as appealing in the valence task, respectively.

Trials for which subjects could not imagine the referent of the noun in the presented color (indicated by pressing a separate button) or for which no response was given were excluded from all analyses.

Experiment 2 differed from Experiment 1 only in that subjects performed the encoding task in the MRI machine and were prompted to make their decision 3 s after word/color onset in order to ensure task processing during the entire trial period. The encoding sessions of Experiments 1 and 2 were followed by a 3-step surprise recognition memory test (administered ~30 min after the encoding session) that was unscanned and self-paced (Fig. 1B). Seven hundred items total were shown (420 old, 280 new), and subjects were first asked to decide if each presented item was old or new. The sequence of old and new test items was pseudorandomly intermixed, so that no more than 3 old or new items were shown successively. If the response was "old," relational/source memory was assessed by querying memory for 1) the color that was presented with the item and 2) the task context in which the item was encountered. Question mark responses were allowed to avoid forced choice guesses. For memory-based fMRI analyses, encoding trials were sorted based on the performance during the retrieval session, yielding the following 5 trial types: 1) items later forgotten (misses), 2) items later recognized, without remembering the correct color or the correct encoding task ("item-only" recognition), 3) items later recognized, including memory for the correct color but not for the encoding task (item and color recognition), 4) items later recognized, including memory for the correct encoding task but not for the color (item and task recognition), and finally 5) items later recognized, including memory for both the correct color and the correct encoding task (item and color and task recognition).

Experiment 3

In order to disentangle the effects of 1) giving an affirmative response (saying yes) and 2) perceived event congruency, we had subjects indicate whether a given word/color combination is "plausible" in one task and whether it is "unusual" in another task. For example, the combination ELEPHANT/red would probably evoke a nonaffirmative response in the plausibility task (no, implausible), whereas it would probably evoke an affirmative response in the unusual task (yes, unusual), although it is perceived as a semantically incongruous combination in both cases. Thus, event congruency was again defined by subjects themselves by rating trials as "plausible" in the plausibility task and as "usual" in the unusual task, respectively.

These tasks were presented in 8 alternating blocks of 36 trials (4.5-s-long trials (see above), separated by a 1-s-long fixation cross), with the first task being counterbalanced across subjects. Encoding was again followed by a surprise recognition memory test (including 144 new words), querying both item memory as well as memory for the associated color (blue, green, red, or yellow).

fMRI Procedures and Analyses

Scanning was performed on a 3T Siemens Allegra MRI system using a whole-head coil. Functional data were acquired using a gradient-echo echo-planar pulse sequence (time repetition = 2.25 s, time echo = 30 ms, 40 slices oriented perpendicular to the hippocampal axis, 3 × 3 × 3 mm voxel size, 0.6-mm interslice gap, 256 volume acquisitions per run). High-resolution T_1-weighted (magnetization-prepared rapid-acquisition gradient echo) images were collected for anatomical visualization. A vacuum pillow minimized head motion. Visual stimuli were projected onto a screen that was viewed through a mirror, and responses were collected with a magnet-compatible button box.

Data were analyzed using SPM2 (Wellcome Department of Cognitive Neurology, London). During preprocessing, images were corrected for differences in slice acquisition timing, followed by motion correction across all runs. Structural and functional data were spatially normalized to an EPI template and voxels were spatially smoothed with a 6-mm FWHM isotropic Gaussian kernel. Statistical analyses were performed using the general linear model implemented in SPM2, entering conditions of interest as regressors of delta functions convolved with a canonical hemodynamic response function and its first-order temporal derivative. Parameter estimates (beta weights) for each regressor of interest were derived for each subject and carried forward to second-level group analyses. All voxel coordinates are reported in Montreal Neurological Institute (MNI) space.

Results

Experiment 1

Experiment 1 yielded 2 main findings. First, the cSME (Note that in this experiment, the factors Congruency and "Response type" are confounded, as congruous events always entail yes responses (see also Schulman 1974 and Craik and Tulving 1975). We choose the term "congruency" instead of "response type/congruency" for brevity and in anticipation of the results from Experiment 3 that show that the cSME is indeed due to the perceived event congruency rather than to saying yes.) was observable not only for the plausibility task that required subjects to rate the perceived semantic match of imagined object/color combinations (plausible or implausible) but also for the valence task that required subjects to rate the perceived aesthetic match of imagined object/color combinations (appealing or unappealing). This is the first report of perceived event congruency enhancing subsequent memory performance in the context of subjective aesthetic judgments. Second, a cSME was seen not only for the target word and the associated color, that is, the constituent elements subject to the congruency judgment but also extended to include enhanced memory for the associated task context (i.e., plausibility or valence judgment made during encoding). This importantly suggests that congruency boosts subsequent memory for the entire encoding event (including the specific encoding task context), not just for the elements whose congruency is evaluated (word and the associated color in our paradigm). These results are summarized in Table 1. Statistical analyses were conducted via repeated-measures analyses of variance (ANOVAs) including the factors Task (plausibility, valence) and Congruency (Congruous, incongruous). Expressed by a main effect of congruency, memory performance was significantly better for congruous events with regard to item memory (hits) ($F_{1,7} = 10.370$, $P = 0.015$, $\eta^2 = 0.597$), color memory ($F_{1,7} = 31.786$, $P = 0.001$, $\eta^2 = 0.820$) and task memory ($F_{1,7} = 28.431$, $P = 0.001$, $\eta^2 = 0.802$). Memory performance did not differ between the 2 encoding Tasks nor was there a Task × Congruency interaction (all $Fs_{1,7} < 0.983$, $P > 0.354$).

An important question with regard to possible mechanisms underlying the cSME was whether congruous events entail longer response times (RTs) than incongruous events, which might explain, at least in part, why those events produce better subsequent memory performance. Average RTs are shown in Table 1, and the statistical analysis revealed no main effects of task, congruency, or a Task × Congruency interaction (all $Fs_{1,7} < 1.047$, $P > 0.340$). The absence of RT differences (see also Craik and Tulving 1975) suggests that the cSME is not due to more time spent on task during congruous encoding events.

Experiment 2 (Functional Brain Imaging)

Note that because the classification of encoding trials as congruous or incongruous was entirely subject dependent in both encoding tasks, the number of trials in each condition of interest (plausible, implausible, appealing, and unappealing trials) could not be controlled a priori via counterbalancing. We therefore derived the ratio of congruous to incongruous trials within each encoding task and statistically tested for potential response biases across subjects via a repeated-measures ANOVA including the factors Task (plausibility, valence) and Congruency (congruous, incongruous). There were no main effects of congruency nor Task × Congruency interactions (all $Fs_{1,21} < 1.822$, $P > 0.191$). The corresponding numbers averaged across subjects were 51.30% "plausible" and 48.70% "implausible" responses in the plausibility task (standard deviation [SD] = 13.69), and 47.60% "appealing" and 52.40% "unappealing" responses in the valence task (SD = 12.33).

With regard to SMEs, all behavioral findings from Experiment 1 were replicated: a main effect of Congruency was found for item memory ($F_{1,21} = 9.209$, $P = 0.006$, $\eta^2 = 0.305$), color memory ($F_{1,21} = 129.101$, $P < 0.001$, $\eta^2 = 0.860$), and task memory ($F_{1,21} = 8.532$, $P = 0.008$, $\eta^2 = 0.289$), all due to congruous events yielding superior subsequent memory compared with incongruous events (Table 2). Again, there

Table 1
Experiment 1—RTs and memory performance

Encoding task	Congruency	RT (s)	% Item memory	% Color memory			% Task memory		
				Correct	Incorrect	"?" Response	Correct	Incorrect	"?" Response
Plausibility	Congruous (plausible)	2.32 (0.17)	76.01 (14.06)	65.07 (17.47)	5.97 (5.70)	28.95 (17.86)	41.66 (17.51)	12.51 (8.75)	45.83 (19.40)
	Incongruous (implausible)	2.35 (0.36)	70.09 (16.06)	42.38 (15.44)	11.55 (8.72)	46.07 (16.68)	28.94 (17.93)	11.41 (8.83)	59.65 (23.21)
Valence	Congruous (appealing)	2.41 (0.37)	77.08 (13.17)	67.21 (15.73)	10.03 (8.93)	22.76 (11.86)	43.28 (17.55)	15.13 (8.33)	41.59 (15.38)
	Incongruous (unappealing)	2.38 (0.39)	69.81 (14.19)	44.87 (17.80)	10.33 (8.19)	44.81 (16.42)	26.19 (15.92)	16.21 (11.12)	57.59 (21.30)

Note: Data represent average values across subjects, with SDs shown in parentheses. Item memory indicates the proportion of items later recognized (hits, as opposed to misses) of all encoding trials belonging to a given event type. Color and task memory are expressed as proportions of all recognized items.

Table 2
Experiment 2—memory performance

Encoding task	Congruency	% Item memory	% Color memory			% Task memory		
			Correct	Incorrect	"?" Response	Correct	Incorrect	"?" Response
Plausibility	Congruous (plausible)	84.95 (12.00)	74.84 (11.98)	12.16 (9.70)	13.00 (9.64)	59.16 (18.90)	18.88 (12.63)	21.97 (15.45)
	Incongruous (implausible)	81.42 (13.12)	62.00 (15.01)	17.60 (12.57)	20.41 (14.20)	56.17 (21.40)	15.93 (11.42)	27.89 (17.58)
Valence	Congruous (appealing)	84.76 (11.46)	74.31 (15.56)	11.17 (8.94)	14.52 (10.73)	60.92 (11.75)	17.27 (8.90)	21.81 (14.32)
	Incongruous (unappealing)	81.47 (11.86)	63.35 (14.24)	15.43 (9.53)	21.22 (14.21)	51.21 (16.92)	18.59 (6.57)	30.20 (19.66)

Note: Data represent average values across subjects, with SDs shown in parentheses. Item memory indicates the proportion of items later recognized (hits, as opposed to misses) of all encoding trials belonging to a given event type. Color and task memory are expressed as proportions of all recognized items.

were no main effects of task nor Task × Congruency interactions (all $Fs_{1,21} < 2.529$, $P > 0.126$).

The first critical question regarding our imaging data was whether there are any brain regions that show enhanced activation for congruous relative to incongruous events while globally supporting successful memory formation in our paradigm. In particular, brain regions underlying the cSME should 1) show enhanced engagement during congruous relative to incongruous events, irrespective of the encoding task and 2) be predictive of subsequent memory performance, irrespective of event congruency. Moreover, considering the notion that the cSME is driven by semantic elaboration, the resulting regions should overlap with regions known to be involved in semantic processing. Thus, we first derived parameter estimates (see Materials and Methods) for congruous and incongruous events, separately for the plausibility and the valence task. Second, we derived parameter estimates for trials yielding relatively low subsequent memory performance (misses or item-only trials) and those yielding relatively high subsequent memory performance (item and color, item and task, or item, color and task trials) separately for congruous- and incongruous encoding events. Finally, in order to identify regions that meet the criteria of 1) showing enhanced activation for congruous versus incongruous events, irrespective of the encoding task and 2) showing enhanced activation for successful versus unsuccessful memory formation, irrespective of event congruency, we conducted a whole-brain quadruple conjunction analysis (Nichols et al. 2005) combining the effect contrasts (a) plausible > implausible, (b) appealing > unappealing, (c) high subsequent memory for congruous events > low subsequent memory for congruous events, and (d) high subsequent memory for incongruous events > low subsequent memory for incongruous events. Note that the conjunction of contrasts (a) and (b) reveals regions more engaged during congruous relative to incongruous events, irrespective of the encoding task, whereas the conjunction of contrasts (c) and (d) reveals regions more engaged during successful relative to unsuccessful memory formation, irrespective of event congruency. Each contrast was thresholded at $P < 0.05$, so that the conjoint significance (in the sense of the logical "and" operator) was assessed at the conservative criterion of at least 10 contiguous voxels exceeding $P < 0.00000625$ (uncorrected).

Three significant clusters emerged from this analysis, located in the anterior–ventral left inferior frontal gyrus (avLIFG), middorsal left inferior frontal gyrus (mdLIFG), and left inferior temporal gyrus (ITG) (Fig. 2 and Table 3). As shown in Figure 2, all 3 regions showed significantly greater activation during congruous compared with incongruous events (all $Ts(21) > 3.45$, $P < 0.003$) and during successful compared with unsuccessful memory formation (all $T_{21} > 4.68$, $P < 0.001$), which is to be expected given the contrasts entered in the underlying conjunction analysis. However, to ensure that the resulting regions did not predict subsequent memory differentially for congruous or incongruous events (i.e., showed different effect sizes depending on event congruency), we conducted repeated-measures ANOVAs with the factors Subsequent Memory (low, high) and Event Congruency (congruous, incongruous) on the corresponding subject-specific beta weights averaged across the resulting clusters. As expected, all 3 regions showed a main effect of subsequent memory (all $Fs_{1,21} > 22.02$, $P < 0.001$) but critically no Subsequent Memory

× Event Congruency interaction (all $Fs_{1,21} < 0.75$, $P > 0.39$). The same type of analysis was conducted to ensure that the effect of event congruency was not greater for one encoding task than the other, employing the factors event congruency (congruous, incongruous) and Task (plausibility, valence). Again, all 3 regions showed the expected main effect of Event Congruency (all $Fs_{1,21} > 12.00$, $P < 0.003$), but critically no Event Congruency × Task interaction (all $Fs_{1,21} < 0.65$, $P > 0.42$). In sum, activation in all 3 regions emerging from our conjunction analysis was enhanced for congruous relative to incongruous events irrespective of the encoding task and at the same time predictive of subsequent memory performance irrespective of event congruency.

Crucially, neuroimaging data have consistently implicated regions along the LIFG and the ITG, overlapping with the ones reported here, in semantic retrieval and elaboration (Thompson-Schill et al. 1997; Wagner et al. 1998; Poldrack et al. 1999; Martin and Chao 2001; Wagner et al. 2001; Gold and Buckner 2002; Badre et al. 2005; Gold et al. 2005; Wig et al. 2005; for recent review, see Martin 2007). This lends empirical support to the notion that the cSME is driven by enhanced semantic elaboration during congruous relative to incongruous events as proposed by Schulman (1974) and Craik and Tulving (1975).

The regions revealed in the above group-level analysis meet the criteria of underlying the cSME by being more engaged during congruous relative to incongruous events and by showing global SMEs. However, our relatively large sample size of 22 subjects allowed us to go one step further and establish a more direct link between brain activity and enhanced memory for congruous encoding events by querying whether there are any brain regions that show differential activation between congruous relative to incongruous events that is correlated with each individual subject's corresponding behavioral cSME. That is, not every subject showed the same boost in subsequent memory performance for congruous relative to incongruous events (the size of the cSME varied from subject to subject) and a critical question is whether there is a brain/behavior correlation between 1) the extent to which specific brain regions are differentially engaged during congruous relative to incongruous events and 2) the size of the cSME. In order to address this question, we first derived a behavioral "benefit index" (BI), which captures the size of the cSME for each subject:

$$BI = \text{proportion of congruous ICT trials}$$
$$- \text{proportion of incongruous ICT trials,}$$

where "ICT" refers to trials that include subsequent item memory along with memory for the associated color as well as for the associated encoding task. The resulting BI for each subject was then, on a voxel-by-voxel level, correlated with differences in brain activation between congruous and incongruous events across subjects. Statistical significance was assessed by applying a threshold of at least 10 contiguous voxels exceeding $P < 0.001$ (uncorrected). Interestingly, we again observed 2 clusters located in anterior–ventral and middorsal portions of the left inferior frontal gyrus that showed strong correlation values (r values of 0.72 in the avLIFG and of 0.79 in the mdLIFG; shown in Fig. 3; complete list of clusters shown in Table 4). Critically, this subject-specific analysis additionally revealed a strong brain/behavior correlation ($r = 0.73$) in the right hippocampus (Fig. 3). The role of the
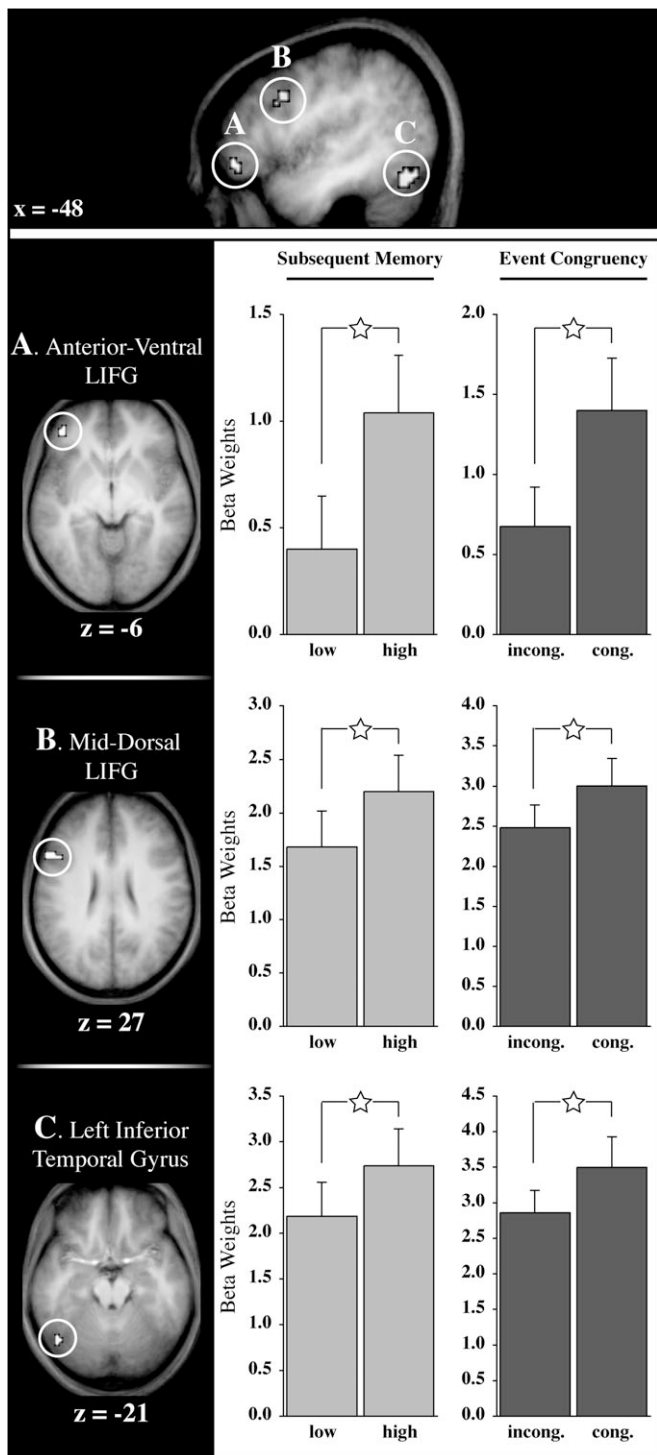
Figure 2. Brain regions underlying the cSME. Statistical parametric maps depict regions that show greater encoding activation for (*A*) plausible compared with implausible trials, (*B*) appealing compared with unappealing trials, (*C*) successful compared with unsuccessful memory encoding during congruous events, and (*D*) successful compared with unsuccessful memory encoding during incongruous events, revealed via a quadruple conjunction analysis. Top: Resulting clusters displayed on a sagittal slice of the T_1-weighted anatomical image averaged across subjects. Bottom: Individual clusters superimposed on horizontal slices of the T_1-weighted anatomical image averaged across subjects. Bar graphs show the beta parameter estimates for each trial type averaged across subjects and across all voxels in a given cluster. Error bars represent the standard error of the mean. low, encoding trials later forgotten (misses) and encoding trials later recognized without memory for the correct color or the correct encoding task (item only recognition); high, encoding trials

Table 3
Brain regions showing greater activation during congruous relative to incongruous encoding events and during successful relative to unsuccessful memory formation, revealed via a quadruple conjunction analysis

Region	Cluster size	Peak *t* value	Peak *x, y, z* coordinates	~BA
L anterior–ventral Inferior frontal gyrus	15	2.57	−48, 45, −12	47
L inferior temporal gyrus	26	2.56	−45, −66, −18	37
L middorsal inferior frontal gyrus	29	2.42	−54, 18, 30	44/9

Note: L, left; Cluster size, number of contiguous suprathreshold voxels; ~BA, approximate Brodmann's areas. Voxel coordinates are reported in Montreal Neurological Institute space.

hippocampus in associative/relational binding has been firmly established (Cohen et al. 1999; Davachi et al. 2003; Ranganath et al. 2004; Staresina and Davachi 2006, 2008; for reviews, see Squire et al. 2004; Davachi 2006; Eichenbaum et al. 2007; Mayes et al. 2007). Tracking the extent to which differential encoding activation during congruous events relates to an individual's behavioral mnemonic benefit for these events, operations supported by these brain regions may ultimately underlie the cSME. In sum, our imaging results suggest that the cSME is driven by enhanced semantic elaboration mediated by increased LIFG engagement in combination with relational binding operations supported by the hippocampus.

Experiment 3

Finally, we asked whether the cSME in our studies was driven by the fact that subjects gave affirmative (yes) responses to congruous word/color combinations or by the fact that those events reflected perceived congruency of the constituent elements. In order to adjudicate between these 2 factors, we used 2 encoding tasks: one that was identical to the plausibility task in our previous studies (a plausibility trial was cued by the question "plausible?") and another where trials were cued by the question "unusual?". If the observed memory boost for congruous events were simply driven by "yes" responses, one would expect events that subjects deem unusual (incongruous/yes) in the unusual task to be remembered as well as plausible (congruous/yes) trials in the plausibility task. Conversely, if the cSME were driven by the perceived event congruency, one would expect events that subjects deem usual (congruous/no) in the unusual task to be remembered as well as plausible (congruous/yes) trials in the plausibility task.

Memory performance (successful item and color memory) was again analyzed via repeated-measures ANOVAs, employing the factors Response (yes, no) and Congruency (congruous, incongruous). We observed a main effect of Congruency on the proportion of both item memory ($F_{1,8} = 14.424$, $P = 0.005$, $\eta^2 = 0.643$) and color memory ($F_{1,8} = 18.464$, $P = 0.003$, $\eta^2 = 0.698$) in the absence of a Response main effect or a Response × Congruency interaction (all $Fs_{1,8} < 3.466$, $P > 0.099$). As shown in Table 5, plausible (congruous/"yes") events yielded better memory than implausible (incongruous/"no") events, whereas usual (congruous/"no") events yielded better memory than unusual (incongruous/"yes") events. This finding suggests that

later recognized including memory for the correct color, the correct task or both; incong., incongruous encoding trials, entailing implausible (plausibility task) or unappealing (valence task) responses; cong., congruous encoding trials, entailing plausible (plausibility task) or appealing (valence task) responses; star symbols, pairwise comparison (*t* test) statistically significant at $P<0.005$.
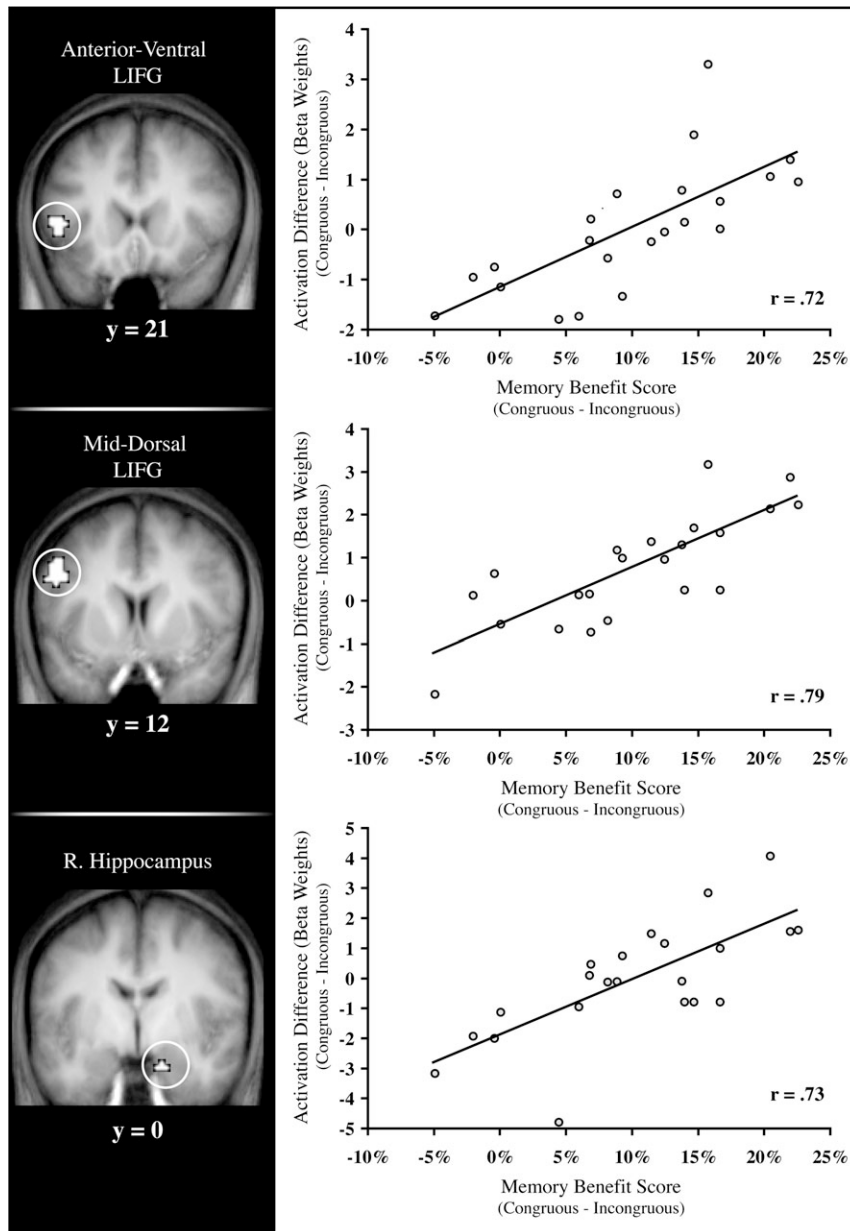
Figure 3. Correlation between differential brain activation and the behavioral memory benefit for congruous events across subjects. Statistical parametric maps (superimposed on coronal slices of the T_1-weighted anatomical image averaged across subjects) depict regions that show significant mapwise correlations between 1) differential brain activation between congruous and incongruous encoding events and 2) the behavioral mnemonic benefit for congruous relative to incongruous events.

the mnemonic advantage of "plausible" and "appealing" trials in our previous studies was driven by the perceived congruency of word/color combinations rather than by eliciting "yes" responses (see also Marks et al. 1992).

Discussion

Our current results suggest that event congruency promotes successful memory formation through enhanced semantic elaboration and relational binding operations. In their attempt to explain the impact of congruency on memory formation, both Schulman (1974) and Craik and Tulving (1975) speculated that congruous events lead to enhanced semantic elaboration as those events can be more readily integrated into a preexisting cognitive

network and the abundant semantic associations will in turn enhance the resulting memory trace. Despite making intuitive sense, however, this account has thus far received no empirical support and alternative theories that emphasize retrieval-based advantages for event congruency have been proposed (Moscovitch and Craik 1976). In the current fMRI study, we first demonstrate that congruency leads to enhanced encoding activation in the avLIFG, the mdLIFG, and the left ITG, all of which were also, and irrespective of event congruency, predictive of successful episodic encoding (Fig. 2). These regions are the ones most consistently associated with semantic retrieval and semantic elaboration throughout the neuroimaging literature (Thompson-Schill et al. 1997; Wagner et al. 1998; Poldrack et al. 1999; Martin and Chao 2001; Wagner et al. 2001; Gold and

Buckner 2002; Badre et al. 2005; Gold et al. 2005; Wig et al. 2005; for recent review, see Martin 2007). Critically, this result offers empirical support for the idea that event congruency may indeed bolster subsequent memory performance through enhanced semantic elaboration. Moreover, congruous trials were accompanied not only by better memory for the target items and their colors (the features directly involved in subjects' judgments) but also for additional event details such as the encoding task in which the item was encountered. Consistent with the effect of semantic elaboration on source memory, this result provides the first evidence that event congruency benefits not only item encoding but also episodic encoding of the entire event, including both perceptual (color) and cognitive (encoding task) source details (Johnson et al. 1993).

However, how exactly does enhanced semantic elaboration map onto superior subsequent memory for all aspects of the encoding event? In order for generated associations to effectively enhance subsequent memory, they must be integrated with the current item representation to form an enriched mnemonic trace. Applying a targeted brain/behavior correlation analysis across our 22 subjects, we found that the size of the cSME, that is, the extent to which event congruency benefits subsequent memory not only correlates with activation increases in both LIFG regions but also in the right hippocampus (Fig. 3). Encoding activation in the hippocampus has consistently been linked to relational memory formation (Cohen et al. 1999; Davachi and Wagner 2002; Davachi et al. 2003; Ranganath et al. 2004; Staresina and Davachi 2006, 2008; for reviews, see Squire et al. 2004; Davachi 2006; Eichenbaum et al. 2007; Mayes et al. 2007). This may suggest that the beneficial effect of event congruency on memory formation ultimately relies on an enhancement of relational binding mechanisms employed during those events. The proposed mechanism of integrating semantic associations via relational binding conforms well with

the finding that the behavioral cSME has been most pronounced for relational memory (cued recall) in the early reports of Schulman (1974) and Craik and Tulving (1975) as well as in our current experiments (memory for the associated color and the associated encoding task; Tables 1 and 2). In sum, our fMRI analyses suggest that event congruency indeed brings online an enriched semantic network associated with the current event. Moreover, we find that the extent to which this semantic network benefits subsequent memory performance relies on additional engagement of the hippocampus, presumably in the service of effectively binding the associations to form a rich episodic memory trace.

Finally, the cSME was evident not only following the plausibility task that explicitly queried the perceived semantic congruency of the constituent elements (the referent of a word and an associated color) but also following the valence task where congruency denoted the perceived aesthetic match of the constituent elements. This important finding highlights that the cSME is present for a match between a stimulus and, perhaps, any internal schema and is reminiscent of other well-documented memory effects, such as the self-referential memory effect (Rogers et al. 1977; Symons and Johnson 1997). However, one potential caveat is that subjects may have based their valence judgments on a semantic match between word and color (i.e., only word/color combinations that are also perceived as plausible are given "appealing" responses), raising the possibility that the beneficial memory effects in the valence task may in fact be mediated by semantic congruency. Although it is clear that our 2 encoding tasks share certain cognitive operations (e.g., word/color processing, imagery, and decision making), 3 sets of results importantly suggest that these 2 tasks indeed differed along critical dimensions. First, source memory for the task performed (plausibility or valence) was high for both tasks (participants from Experiment 2 only chose the incorrect source 18% of the time [SD = 7]). If the 2 tasks were indistinguishable in terms of the encoding operations employed, the resulting memory trace would not be distinctive enough to allow for source memory scores in that range. Second, as shown in the Supplementary Material, we compared fMRI encoding activation between the plausibility task and the valence task and found that the valence task differentially activated, among other regions (Supplementary Table S1), large areas in medial prefrontal cortex (Supplementary Fig. S1). The consistent involvement of this region in self-referential processing (for review, see Amodio and Frith 2006) nicely aligns with the intended emphasis of the valence task on introspective aesthetic decisions and provides additional neuroimaging evidence for differential cognitive processes employed during our 2 encoding tasks. Finally, as

Table 4

Regions emerging from the brain/behavior correlation between 1) differential encoding activation and 2) the corresponding behavioral mnemonic benefit for congruous relative to incongruous events

Region	Cluster size	Peak t value	Peak x, y, z coordinates	~BA
L middorsal inferior frontal gyrus	64	5.74	−51, 9, 24	44/9
R hippocampus	11	5.68	15, 3, −27	—
R superior parietal lobule	15	5.27	30, −60, 63	7
L medial superior frontal gyrus	25	4.79	−3, 24, 51	8
L anterior–ventral inferior frontal gyrus	24	4.59	−48, 21, 3	47/45

Note: L, left; R, right; Cluster size, number of contiguous suprathreshold voxels; ~BA, approximate Brodmann's areas. Voxel coordinates are reported in Montreal Neurological Institute space.

Table 5

Experiment 3—memory performance

Encoding question	Response	% Item memory	% Color memory		
			Correct	Incorrect	"?" Response
Plausible?	Yes (plausible/congruous)	87.80 (6.03)	74.25 (16.47)	13.70 (15.02)	12.05 (11.88)
	No (implausible/incongruous)	76.88 (12.99)	61.76 (20.36)	17.71 (16.62)	20.52 (18.70)
Unusual?	No (usual/congruous)	86.27 (9.45)	78.88 (14.55)	10.02 (14.45)	11.10 (11.41)
	Yes (unusual/incongruous)	82.41 (13.22)	63.05 (17.20)	19.70 (16.86)	17.25 (16.45)

Note: Data represent average values across subjects, with SDs shown in parentheses. Item memory indicates the proportion of items later recognized (hits, as opposed to misses) of all encoding trials belonging to a given event type. Color memory is expressed as proportions of all recognized items.

described in detail in the Supplementary Material, we conducted a normative study with an additional 20 subjects in which we derived plausibility ratings for each possible word/color combination used in our current studies. For each word, the plausibility of all 4 colors (blue, green, red, and yellow) was rated on a 4-point scale (plausible high, plausible medium, implausible medium, and implausible high). We then evaluated the match between 1) those independent normative ratings and 2) the valence and plausibility judgments across subjects in our current experiments. For both Experiments 1 and 2, we found that the valence task resulted in a significantly reduced match with the normative ratings than the plausibility task, whereas no differences were seen between the plausibility task and the unusual task in Experiment 3. In other words, word/color combinations that were rated as plausible in the normative study were less likely to be rated as appealing than as plausible in Experiments 1 and 2. Taken together, these results strongly suggest that the plausibility and the valence task in our experiments were not entirely overlapping. Although certainly more experimental work will be needed to systematically assess the extent to which perceived event congruency can be independent from a purely semantic match of the constituent elements, our current findings open the possibility that the cSME may apply to a broader definition of event congruency than initially revealed by the seminal papers of Schulman (1974) and Craik and Tulving (1975).

Taken together, the finding that event congruency bolsters episodic encoding sheds an interesting light on the relationship between episodic and semantic memory. The impact of semantic memory on episodic encoding has been discussed in theoretical accounts such as the "Serial Parallel Independent" model proposed by Tulving and Markowitsch (1998), where the authors emphasize that semantic knowledge is a necessary precursor to episodic memory formation, that is, information must be encoded into episodic memory "through" semantic memory. Moreover, in a series of elegant behavioral experiments, Bransford and Johnson (1972, 1973) demonstrated that being provided with a congruous context before hearing a short narrative largely enhances subsequent recall for various elements of the narrative. Critically, there was no beneficial effect when the context was provided after hearing the narrative or when an incongruous context was provided before hearing the narrative. In other words, being able to embed new incoming information into an existing cognitive schema (i.e., the presence of perceived event congruency) had a strong positive effect on successful episodic encoding. Finally, building on these early behavioral findings, the impact of existing conceptual knowledge on episodic encoding has recently received much attention when Tse et al. (2007) reported that in rats, new information that is congruous with a previously established associative schema/concept is more effectively encoded, that is, requires a shorter hippocampal consolidation period after single-trial learning. The cSME reported in our current paper nicely aligns with these theoretical accounts, behavioral findings and recent findings from animal studies, highlighting the close mnemonic interdependence of existing cognitive concepts and novel episodic encoding.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/

References

Amodio DM, Frith CD. 2006. Meeting of minds: the medial frontal cortex and social cognition. Nat Rev Neurosci. 7:268–277.

Badre D, Poldrack RA, Paré-Blagoev EJ, Insler RZ, Wagner AD. 2005. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. Neuron. 47:907–918.

Bransford JD, Johnson MK. 1972. Contextual prerequisites for understanding: some investigation of comprehension and recall. J Verb Learn Verb Behav. 11:717–766.

Bransford JD, Johnson MK. 1973. Considerations of some problems of comprehension. In: Chase W, editor. Visual information processing. New York: Academic Press. p. 383–438.

Cohen NJ, Eichenbaum H. 1993. Memory, amnesia, and the hippocampal system. Cambridge: MIT Press.

Cohen NJ, Ryan J, Hunt C, Romine L, Wszalek T, Nash C. 1999. Hippocampal system and declarative (relational) memory: summarizing the data from functional neuroimaging studies. Hippocampus. 9:83–98.

Craik FIM, Lockhart RS. 1972. Levels of processing: a framework for memory research. J Verb Learn Verb Behav. 11:671–684.

Craik FIM, Tulving E. 1975. Depth of processing and the retention of words in episodic memory. J Exp Psychol Gen. 104:268–294.

Dale AM. 1999. Optimal experimental design for event-related fMRI. Hum Brain Mapp. 8:109–114.

Davachi L. 2006. Item, context and relational episodic encoding in humans. Curr Opin Neurobiol. 16:693–700.

Davachi L, Mitchell JP, Wagner AD. 2003. Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. Proc Natl Acad Sci USA. 100:2157–2162.

Davachi L, Wagner AD. 2002. Hippocampal contributions to episodic encoding: insights from relational and item-based learning. J Neurophysiol. 88:982–990.

Demb JB, Desmond JE, Wagner AD, Vaidya CJ, Glover GH, Gabrieli JD. 1995. Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional MRI study of task difficulty and process specificity. J Neurosci. 15:5870–5878.

Eichenbaum H, Yonelinas AP, Ranganath C. 2007. The medial temporal lobe and recognition memory. Annu Rev Neurosci. 30:123–152.

Gold BT, Balota DA, Kirchhoff BA, Buckner RL. 2005. Common and dissociable activation patterns associated with controlled semantic and phonological processing: evidence from fMRI adaptation. Cereb Cortex. 15:1438–1450.

Gold BT, Buckner RL. 2002. Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks. Neuron. 35:803–812.

Hall DM, Geis MF. 1980. Congruity and elaboration in free and cued recall. J Exp Psychol [Hum Learn]. 6:778–784.

Johnson MK, Hashtroudi S, Lindsay DS. 1993. Source monitoring. Psychol Bull. 114:3–28.

Lockhart RS, Craik FIM. 1990. Levels of processing: a retrospective commentary on a framework for memory research. Can J Psychol. 44:87–112.

Marks W, McFalls EL, Hopkinson PL. 1992. Encoding pictures in scene contexts: does task demand influence effects of encoding congruity? J Exp Psychol Learn Mem Cogn. 18:192–198.

Martin A. 2007. The representation of object concepts in the brain. Annu Rev Psychol. 58:25–45.

Martin A, Chao LL. 2001. Semantic memory and the brain: structure and processes. Curr Opin Neurobiol. 11:194–201.

Mayes A, Montaldi D, Migo E. 2007. Associative memory and the medial temporal lobes. Trends Cogn Sci. 11:126-135.

Moscovitch M, Craik FIM. 1976. Depth of processing, retrieval cues and uniqueness of encoding as factors in recall. J Verb Learn Verb Behav. 15:447-458.

Nichols T, Brett M, Andersson J, Wager T, Poline JB. 2005. Valid conjunction inference with the minimum statistic. Neuroimage. 25:653-660.

Otten LJ, Henson RNA, Rugg MD. 2001. Depth of processing effects on neural correlates of memory encoding. Brain. 124:399-412.

Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JDE. 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage. 10:15-35.

Ranganath C, Yonelinas AP, Cohen MX, Dy CJ, Tom SM, D'Esposito M. 2004. Dissociable correlates of recollection and familiarity within the medial temporal lobes. Neuropsychologia. 42:2-13.

Rogers TB, Kuiper NA, Kirker WS. 1977. Self-reference and the encoding of personal information. J Pers Soc Psychol. 35:677-688.

Schulman AI. 1974. Memory for words recently classified. Mem Cognit. 2:47-52.

Squire LR, Stark CEL, Clark RE. 2004. The medial temporal lobe. Annu Rev Neurosci. 27:279-306.

Staresina BP, Davachi L. 2006. Differential encoding mechanisms for subsequent associative recognition and free recall. J Neurosci. 26:9162-9172.

Staresina BP, Davachi L. 2008. Selective and shared contributions of the hippocampus and perirhinal cortex to episodic item and associative encoding. J Cogn Neurosci. 20:1478-1489.

Stark CEL, Squire LR. 2001. When zero is not zero: the problem of ambiguous baseline conditions in fMRI. Proc Natl Acad Sci USA. 98:12760-12766.

Symons CS, Johnson BT. 1997. The self-reference effect in memory: a meta-analysis. Psychol Bull. 121:371-394.

Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proc Natl Acad Sci USA. 94:14792-14797.

Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, Witter MP, Morris RGM. 2007. Schemas and memory consolidation. Science. 316:76-82.

Tulving E, Markowitsch HJ. 1998. Episodic and declarative memory: role of the hippocampus. Hippocampus. 8:198-204.

Wagner AD, Paré-Blagoev EJ, Clark J, Poldrack RA. 2001. Recovering meaning left prefrontal cortex guides controlled semantic retrieval. Neuron. 31:329-338.

Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL. 1998. Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. Science. 281:1188-1191.

Wig GS, Grafton ST, Demos KE, Kelley WM. 2005. Reductions in neural activity underlie behavioral components of repetition priming. Nat Neurosci. 8:1228-1233.