

**PUBLISHED AS:**

Henson, R.N.A. (2004). Explicit memory. Human Brain Function, 2nd Edition. Frackowiak, Friston, Frith, Dolan & Price (Eds.), pp. 487-498. Elsevier, London.

Explicit Memory

R. N. A. Henson

Institute of Cognitive Neuroscience &  
Wellcome Department of Imaging Neuroscience, Institute of Neurology  
University College London

Dr. Richard Henson

Institute of Cognitive Neuroscience

17 Queen Square

London WC1N 3AR U.K.

Tel1: +44 (020) 7679 1131

Tel2: +44 (020) 7833 7483

Fax: +44 (020) 7813 1420

Email: r.henson@ucl.ac.uk

**Acknowledgements**

This work is funded by Wellcome Trust Fellowship 060924. The author would like to thank Mick Rugg and Ray Dolan for many useful discussions.

Total pages (inc legends): 16    Figures: 5    Words (body+refs+legends): 6,885

## Contents

---

<b>I</b>	<b>Introduction</b>
<b>II</b>	<b>Introduction: Recognition Memory</b>
<b>III</b>	<b>Basic Old-New effects</b>
<b>IV</b>	<b>Familiarity vs. Recollection</b>
<b>V</b>	<b>New-Old effects</b>
<b>VI</b>	<b>Subsequent memory effects</b>
<b>VII</b>	<b>Conclusions</b>
	<b>References</b>

---

## I. Introduction

A previous chapter (Henson, this volume) focused on neuroimaging studies of repetition priming, a form of implicit memory. The present chapter concerns imaging studies of recognition memory, a form of explicit memory (see Henson, this volume, for definition of these terms). The haemodynamic correlates of recognition memory are indexed by the same basic repetition effect as are the correlates of repetition priming, namely the difference between repeated (“old”) and initial (“new”) presentations of a stimulus within an experimental context. The difference is that recognition memory is tested by a “direct” memory task, in which participants discriminate old from new stimuli (as distinct from the “indirect” tasks discussed in Henson, this volume). Like the previous chapter, the present chapter is selective in focusing on recognition memory (as distinct from other tests of explicit memory, such as recall), and also in focusing only on studies performed at the Wellcome Department of Imaging Neuroscience (see Buckner & Wheeler, 2001; Desgranges et al., 1998; Fletcher & Henson, 2001; Rugg, 2002, for more comprehensive reviews).

## II. Introduction: Recognition Memory

Whereas repetition priming has typically been associated with a reduced haemodynamic response for repetition versus control conditions (repetition suppression), recognition memory has typically been associated with an increased response for repetition conditions (repetition enhancement, or an “old-new” effect). As with priming however, this generalisation may not always hold (see Section V below). The old-new effect is normally conditionalised on correct recognition memory judgments, namely the difference between “hits” (old stimuli called “old”) and “correct rejections” (new stimuli called “new”), and excluding “misses” (old stimuli called “new”) and “false alarms” (new stimuli called “old”). This is the case for all the event-related studies described below.

Recognition memory has traditionally been conceptualised in terms of signal-detection theory (Green & Swets, 1966), in which stimuli are assumed to have a continuous range of “strengths” in memory. Old and new stimuli are represented by two overlapping distributions along this continuum, with the central tendency of old stimuli being greater than that of new stimuli. The difference in these central tendencies is called the “discriminability”. The participant is assumed to place a decision boundary somewhere on the continuum, called the “response criterion”, which they use to categorise stimuli as either “old” or “new”. The strength of an old or new item relative to this criterion determines whether it becomes a hit, miss, correct rejection or false alarm (Figure 1). Estimates of the discriminability (e.g,  $d'$ ) and response criterion (e.g,  $\beta$ ) can be obtained by various methods, though a summary measure like “pH-pFA”, the proportion of old words that are hits minus the proportion of old words that are false alarms, is often a satisfactory index of memory accuracy.

Though the signal detection model can explain some important results in recognition memory, many researchers have argued that it is insufficient (Mandler, 1980). Consequently, “dual-process” models have been developed that propose two distinct contributions to recognition memory: “recollection” and “familiarity” (both examples of explicit memory). Recollection refers to retrieval of the specific event in the past in which a stimulus was presented (e.g, time and place) and is associated with “episodic memory”. Familiarity on the other hand is associated with a feeling that a stimulus has been experienced before, in the absence of memory

for the context in which that stimulus occurred. Recollection is often viewed as multidimensional and effortful; familiarity is often viewed as a unidimensional quantity and relatively automatic. In some models, familiarity is equated with the memory strength in signal detection theory, whereas recollection is modelled as an independent, all-or-none occurrence (Yonelinas et al., 1996). The precise manner in which familiarity and recollection combine in a recognition memory judgment remains controversial however (Knowlton & Squire, 1995; see Section III below).

The next section covers the old-new effects observed in simple “yes/no” recognition memory tests, including early PET studies. Section IV deals with variants of recognition memory tests, in which decisions are augmented by indications of the subjective experience or confidence accompanying each decision, or which are conditional on a requirement to retrieve some aspect of the study episode (called tests of “source memory”). These variants have been used in attempts to dissociate the neural correlates of recollection and familiarity. Section V discusses new-old effects (or repetition suppression) in recognition memory tests. Finally, Section VI discusses a slightly different topic that does not concern repetition effects per se, but rather studies of memory encoding that examine responses to initial presentations of stimuli that predict subsequent memory when those stimuli are repeated.

### III. Basic Old-new effects

Early PET and fMRI studies of recognition memory, before the advent of event-related fMRI, investigated old-new effects by comparing blocks in which a high proportion of stimuli were old with blocks in which a low proportion were old. For example, the PET study of Rugg et al. (1996) used this method to compare blocks of 80%, 20% and 0% studied words. In a control task, participants distinguished between two symbol strings, with the proportion of “target” strings varying in the same ratios. Regions in which the response increased with the proportion of old items in the memory task, having covaried out the proportion of targets in the control task, included bilateral anterior prefrontal cortex (+38 +48 +8 and -30 +46 -4), right dorsolateral prefrontal cortex (+42 +22 +28) and medial prefrontal cortex (+4 +20 +48). In a similar comparison of 0% and 80% studied words, Rugg et al. (1998a) found old-new effects in right anterior prefrontal (+32 +50 +18), left lateral parietal (-36 -80 +40) and medial parietal cortex (-2 -74 +36).

However, there are a number of reasons why such blocked repetition effects are difficult to interpret. People are known to be sensitive to event probabilities (Fitzgerald & Picton, 1981). Thus, even though attempts can be made to disguise the old:new ratio manipulation (e.g. by having more balanced old:new ratios during “run-in” periods before and/or after each block), participants are likely to notice that they have made one response more often than the other during the last few trials. A low incidence of “old” responses during a 20%-old block, for example, may lead participants to question their memory, and perhaps adjust their response criterion accordingly (to be more lenient). This means that any differences between the mean haemodynamic response during blocks of different ratios may reflect differences in expectancies or strategies, rather than old-new memory effects per se.

These problems are not so acute for event-related fMRI studies, in which old and new stimuli are randomly intermixed (in equal proportions), and the mean stimulus-locked responses compared directly. Indeed, such methods can be used to test directly the interaction between probability effects and old-new effects. Herron et al. (in prep), for example, compared the event-

related repetition effect for old versus new words for three sessions with different old:new ratios (either 75:25, 50:50, or 25:75). Whereas lateral and medial parietal regions showed repetition enhancement that appeared to be independent of the probability of old items, several prefrontal regions showed an interaction between old:new ratio and repetition. Indeed, some of these prefrontal regions showed a switch from repetition enhancement under the low old:new ratio to repetition suppression under the high old:new ratio. In other words, the responses in these regions were greater for the rarer type of item. Though the reason for this cross-over interaction is unclear (it could relate to the fact that participants were informed of the ratio manipulation, which may cause them to treat the rarer items as the “targets”; Wagner et al., 1998a), it provides strong evidence for strategic effects in such designs, and so questions the results from previous blocked studies. The data also promote a general picture of “automatic” retrieval effects in posterior regions and strategic (task-dependent) post-retrieval effects in frontal regions.

Nonetheless, despite the fact that a number of event-related fMRI studies of “yes/no” recognition memory have used unequal old:new proportions, and that some of the associated prefrontal activations may relate to post-retrieval decision processes, a consistent pattern of regions showing old-new effects has emerged (Rugg & Henson, in press). Maratos et al. (2001) for example, found a network of regions showing repetition enhancement, including left anterior prefrontal cortex (-20 +64 +12), left parahippocampal cortex (-16 -28 -8), posterior cingulate (+4 -54 +18), precuneus (-6 -58 +36) and bilateral lateral parietal cortex (-36 -62 +56; +34 -68 +40), even though the old:new ratio was 3:1. Interestingly, this “retrieval success” network was common to three types of old word – negative, positive or neutral – that had been studied in the context of sentences describing emotionally negative, positive or neutral situations respectively. Emotional modulation of repetition effects, which probably reflected episodic retrieval of the associated study sentence, were also observed in regions including left amygdala, right dorsolateral prefrontal cortex and posterior cingulate for negative emotions, and orbitofrontal, right anterior prefrontal and left anterior temporal cortex for positive emotions.

#### **IV. Recollection vs. Familiarity**

Other studies have used variants of the basic yes/no recognition memory task. Henson et al. (1999b), for example, asked participants to make a three-way decision to old and new words, either “Remember”, “Know” or “New”. This “Remember-Know” or “R-K” distinction was proposed by Tulving (1985) to capture the subjective difference between recalling a prior episode (e.g. what occurred before or after a studied item) and a feeling of “oldness” in the absence of memory for a specific occurrence (a phenomenological distinction that might map onto the hypothetical processes of recollection and familiarity). Note that these decisions are assumed to be qualitatively, and not simply quantitatively, different (though see Donaldson, 1996). Thus, while people are normally confident of their R judgments, they can still be highly confident of their K judgments too (Gardiner et al., 1994). In the Henson et al. (1999b) study, though the pH-pFA rate for K judgments was less than that for R judgments, it was still greater than zero (under various scoring assumptions), indicating that K judgments were not mere guesses.

Collapsing across R and K judgments produced old-new effects similar to those described in Section III. Direct comparisons of correct R versus correct K judgments (Figure 2) revealed greater responses for R judgments in regions including left anterior superior prefrontal (-21 +54 +39), lateral parietal (-57 -51 +39) and posterior cingulate (0 -30 +36) regions, and greater

responses for K judgments in regions including right dorsolateral prefrontal cortex (+51 +30 +27), anterior cingulate (-12 +9 +36) and dorsal precuneus (-12 -60 +57). The former results suggest that the anterior prefrontal, lateral parietal and posterior cingulate regions showing old-new effects in simple “yes/no” recognition memory tests represent (processes contingent on) recollection of the study episode.<sup>1</sup> Given their somewhat contextually-impooverished study phases, the source information associated with recollection in these experiments is most likely to be “internal source” (Johnson et al., 1993), such as the semantic associations made to the words presented at study.

According to a dual-process model, the relative increases for K judgments might implicate right prefrontal and anterior cingulate cortices in support of a familiarity process. However, this type of inference depends on assumptions about how the hypothetical constructs of recollection and familiarity relate to R and K judgments, for example, whether they are independent, redundant or exclusive (Knowlton & Squire, 1995). The preferred interpretation of Henson et al. (1999b) was that the right prefrontal and anterior cingulate activations reflect increased post-retrieval “monitoring” (Henson et al., 1999a; Shallice et al., 1994) for K than R judgments. In other words, when a word seems familiar, but one cannot remember the specific episode in which it was studied, one attempts additional verification of the retrieved information before reaching a decision. This is consistent with the observation of longer RTs for K than R judgments in this study. The familiarity signal itself is assumed to arise elsewhere (such as perirhinal cortex; see Section V).

This hypothesis was tested in a subsequent study by Henson et al. (2000), which used confidence judgments, rather than R-K judgments. For each word, participants made a four-way decision of “confident old”, “unconfident old”, “unconfident new” or “confident new”. Appealing to a signal detection model, the authors reasoned that low confidence judgments (which are close to the response criterion, Figure 1) should require greater monitoring than high confidence judgments, regardless of whether those judgments are “old” or “new”. Consistent with this hypothesis, the right dorsolateral prefrontal region identified in their previous study (Henson et al., 1999b) showed a main effect of low versus high confidence, reinforcing its role in post-retrieval monitoring rather than memory per se. Regions showing a main effect of old versus new words included the usual network of left anterior prefrontal (-21 +63 +21), posterior cingulate (+3 -42 +21), precuneus (0 -69 +33) and left lateral parietal cortex (-48 -57 +48). The posterior cingulate/precuneus regions showed a particularly large response to high confidence old judgments, consistent with a role for these regions in recollection (Figure 3).

In another study attempting to isolate recollection, Cansino et al. (2002) used pictures rather than words, presented in one of four spatial quadrants at study. At test, a source memory task was used, in which participants made a five-way decision to indicate whether the pictures were new, or, if they were old, in which spatial location they occurred at study. Over 60% of responses to old words represented a correct source judgment, and so the majority could be associated with recollection, given that the chance rate was 25%. Regions showing increased responses associated with recollection included left anterior prefrontal (-10 +58 +32), right lateral parietal (+62 -34 +24), left parahippocampal (-14 -44 -4) and right hippocampal formation (+26 -16 -14). The left anterior prefrontal finding is consistent with that of Henson et

---

<sup>1</sup> Another study using the R/K paradigm by Eldridge et al. (2000) reported similar regions showing greater responses to R than K judgments, but additional repetition enhancement in the hippocampus (though pH-pFA for K judgments was close to zero, at least under exclusivity or independence scoring assumptions, suggesting that many K judgments were guesses).

al. (1999b), supporting a role of this region in recollection; the right rather than left parietal finding may reflect a different type of source memory (e.g, spatial, rather than verbal associations). The reverse contrast of incorrect versus correct source judgments revealed bilateral dorsolateral prefrontal (+44 +30 +40; -36 -36 +36) and intraparietal sulcus (+28 -56 +46; -24 -66 +40). The dorsolateral finding is consistent with the K-R findings of Henson et al. (1999b), supporting a role for this region in post-retrieval monitoring when source information is elusive.

Rugg et al. (2003) required source memory judgments for words presented in one of two colours and in one of two spatial locations. In this case, an “exclusion” task was used (Jacoby et al., 1993), in which participants pressed one key if a test word was presented in a specific study context (e.g, red and left of fixation) – the so-called “targets” – and another key if a test word was studied in a different context, or was studied but the context forgotten, or was new – the “non-targets”. The data from this task were compared with those from a standard “yes/no” recognition task (or “inclusion” task), in which one key was pressed for old words (regardless of source) and the other key for new words.<sup>2</sup> Analysis was restricted to prefrontal cortex. Regions showing old-new effects common to both the Inclusion and Exclusion tasks included bilateral anterior prefrontal cortex (-45 +48 +6; +39 +60 -3). Note that these old-new effects also comprised greater responses to correctly excluded nontargets than correct rejections in the Exclusion task, even though both conditions were associated with the same key-press (i.e, controlling for target detection or response selection effects; cf. Section III). A region in the right dorsolateral prefrontal cortex (+48 +42 +24) showed a greater (and delayed) response to hits in the Exclusion task than to hits in the Inclusion task. Because the Exclusion but not Inclusion task requires verification of the spatial/colour source associated with old items, this dorsolateral prefrontal response is again consistent with a role for this region in post-retrieval monitoring (in that monitoring can operate over either recollected or familiarity-driven information).

In summary, a number of recent event-related fMRI studies of recognition memory have identified a common set of regions showing basic old-new effects (repetition enhancement). These include regions of left and right anterior and dorsolateral prefrontal cortex, and in lateral and medial parietal cortices. Furthermore, some of these regions appear to covary with recollection (as operationalised by R judgments, high confidence judgments or correct source judgments), particularly in left anterior prefrontal, left lateral parietal and posterior cingulate cortices. Responses in other regions, such as dorsolateral prefrontal cortex, seem better explained in terms of post-retrieval decision processes. These responses may therefore vary with the specific type of decision required, or where subjects place their response criterion. An important consequence of the latter is that these regions may show either an old-new or a new-old effect, depending on whether the response criterion is closer to the central tendency of the old or new distribution respectively.

However, these prefrontal and parietal old-new effects are not those that might be expected from neuropsychological studies of amnesia, which strongly implicate medial temporal lobe (MTL) structures in explicit memory. The failure to find MTL old-new effects in many of the above studies (with the exception of parahippocampal old-new effects in Maratos et al., 2001, and Cansino et al., 2002) could simply reflect reduced sensitivity within the MTL. Alternatively, the lack of old-new differences might reflect the involvement of MTL in both the encoding of new items (see Section VI), and the retrieval of old items, such that there is no net repetition

---

<sup>2</sup> In attempt to control for differences in non-specific “difficulty” between the Inclusion and Exclusion tasks, an orthogonal manipulation of study list length was used. This manipulation did not interact with any of the other contrasts.

effect. However, preliminary results (below) suggest that some MTL structures do show repetition effects, but these are new-old effects, rather than old-new effects.

## V. New-old effects

Surprisingly, few of the above studies considered regions that showed a reduced response for old relative to new words (i.e, repetition suppression). In a meta-analysis, Henson et al. (in press) found a region in right anterior MTL (+22 -6 -28), more likely in perirhinal cortex than hippocampus, that showed a “new-old” effect common to four different experiments. The effect was found regardless of whether the stimuli were words or pictures, or whether the memory test was direct or indirect. Furthermore, the effect did not appear sensitive to whether a source memory judgment to the old stimuli was correct or incorrect (in the Cansino et al., 2002, study), or whether the task did or did not require source retrieval (in the Rugg et al., 2003, study). These findings suggest that this new-old effect is independent of recollection, but are consistent with a familiarity signal (at least within a redundancy or independence dual-process model).

However, this anterior MTL repetition suppression could also reflect priming (at a high-level of the visual processing pathway) – i.e, an implicit memory effect. This possibility cannot be refuted by the imaging data. Nonetheless, there is strong evidence from lesion (Meunier et al., 1993) and electrophysiological studies (Brown & Xiang, 1998) in animals for a role of perirhinal cortex in explicit recognition memory. Moreover, human lesion studies (Hamann & Squire, 1997) suggest that MTL damage that includes perirhinal cortex does not affect perceptual priming. Thus the anterior MTL repetition suppression is likely to reflect a familiarity signal that is used for explicit recognition memory judgments.

Interestingly, the response to old items in these MTL regions was a deactivation relative to the interstimulus (fixation) baseline. This is not a conceptual problem, since it is possible that repetition of a stimulus produces a sparse pattern of neural firing, in which only a small subset of neurons fire, while the rest are suppressed below spontaneous firing levels (as found in some single-cell studies of human MTL, Fried et al., 2002). More puzzling is the fact that new items did not appear to activate this region above interstimulus baseline. This is certainly a problem for a priming account in which repetition suppression is attributed to the facilitation of a process occurring on first presentations of a stimulus (in which case, one might expect significant response to new items versus baseline; see Section II in Henson, this volume).

The association of anterior MTL cortex with a familiarity signal is consistent with the theory of Aggleton and Brown (1998). These authors used data from human and nonhuman lesions to argue that the recollection and familiarity processes assumed by dual-process models are supported by dissociable structures within MTL: with a hippocampal-anterior thalamic circuit associated with recollection, and a perirhinal-medial dorsal thalamic circuit associated with familiarity. In a post hoc observation, Mick Rugg (personal communication) found a significant interaction in the Cansino et al. (2002) study between correct and incorrect source judgments and two MTL regions (Figure 4). The right anterior cortical region (+22 -6 -28, identified in the meta-analysis of Henson et al. (in press) showed equivalent decreases for correct and incorrect source judgments relative to correct rejections, whereas a slightly more posterior and superior region, most likely in the right hippocampus (+26 -16 -14), showed less of decrease for correct than incorrect source judgments. Though this single dissociation clearly requires replication, it supports the anatomical realisation of recollection and familiarity suggested by



Aggleton and Brown (1998)<sup>3</sup>.

## VI. Subsequent Memory Effects

This final section is not concerned with repetition effects per se, but rather differences between items at study according to whether or not they are later remembered – the so-called “subsequent memory effect”. In this paradigm, participants are scanned at study (rather than test) while performing a task on presented items. After a short delay, they are given a memory test (usually a recognition memory test with concurrent confidence or source judgments), which is used to sort items at study into those remembered and those forgotten. The memory test is normally unexpected (i.e. participants are not aware during study that their memory for the items will be tested later). Note that a similar logic could be used for priming (i.e. to identify regions that predict the size of the subsequent priming effect), as has been done with ERPs (Schott et al., 2002).

Imaging studies of subsequent memory effects generally find greater responses to remembered than forgotten items in ventral prefrontal (e.g.  $-45 +24 -6$  in Henson et al., 1999b;  $-36 +36 -9$  in Otten et al., 2001) and midfusiform cortices (e.g.  $-38 -54 -12$  in Cansino et al., 2002;  $-48 -54 -18$  in Otten et al., 2001), often stronger on the left with words. However, it is important to note that these effects are normally only seen for items subsequently recognised with high confidence (Otten et al., 2001; Wagner et al., 1998b), with R judgments (Brewer et al., 1998; Henson et al., 1999b), or correct source judgments (Cansino et al., 2002). Thus these regions seem specifically to predict subsequent recollection (or episodic retrieval), rather than subsequent familiarity.<sup>4</sup> This is consistent with a study that focussed on the MTL (Strange et al., 2002), in which regions of left hippocampus ( $-22 -26 -16$ ) and left perirhinal cortex ( $-30 -4 -36$ ) predicted subsequent free recall of items from short lists of intentionally-studied words. The relationship between this perirhinal encoding effect at study and the perirhinal familiarity effect at test (in Section V) remains to be determined.

One important question is whether these regions predict subsequent recollection regardless of the study task (as would be the case if they were part of a specialised memory system, the so-called “structuralist” perspective, Schacter & Tulving, 1994), or whether different regions predict recollection under different study and/or test tasks (a “proceduralist” perspective, Kollers & Roediger, 1984). Otten et al. (2001) examined the role of the study task by asking participants to make either semantic or orthographic decisions about words. The semantic decision was whether the word referred to an animate or inanimate entity; the orthographic decision was whether the first and last letters of the word were in alphabetical order. A left ventral prefrontal region ( $-45 +24 -6$ ) and a left anterior hippocampal region ( $-27 -15 -12$ ) predicted subsequent memory under both the semantic and orthographic tasks. This could be taken as support for the structuralist view that there exists a specialised memory system (in the MTL) that, via interactions with ventral prefrontal cortex, allows successful memory regardless of study task.

In a subsequent study however, Otten and Rugg (2001) compared subsequent memory effects under a semantic task with those under a phonological study task. The semantic decisions

---

<sup>3</sup> Though it is less clear why the hippocampal response for correct source judgments, while above that for incorrect source judgments, was still below that for correct rejections of new items.

<sup>4</sup> Note that these effects are unlikely to reflect random variations in attention or arousal at study, given that reaction times to perform the study task do not typically differ for words later remembered and words later forgotten (Otten et al., 2001).

were again animate/inanimate; the phonological decision was whether the word contained an odd or even number of syllables. In this case, a double dissociation was found between left ventral prefrontal (-51 +18 +15) and medial prefrontal (-3 +48 +33) regions that showed greater subsequent memory effects under the semantic than phonological task, and left intraparietal (-39 -45 +51), left fusiform (-42 -48 -15) and left occipital (-24 -75 +33) regions that showed a subsequent memory effect for the phonological but not semantic study task (Figure 5). These data therefore support the proceduralist view that regions predicting memory encoding vary with the study task. In other words, memory is a by-product of the particular processes performed on an item at study.

Thus some support can be found for both the structuralist and the proceduralist views. One possibility is that the components of an episodic memory are stored across different cortical regions (the specific components being those that are emphasized by the study task), but that, in order to bind these components together, they need to be associated with an “index” to that memory stored in the MTL (which represents a specialised memory system).

Another interesting question is whether there are variations in people’s psychological “state” that predict subsequent memory (Rugg & Wilding, 2000), in addition to variations in their responses to individual stimuli. Otten et al. (2002) addressed this question using alternating epochs (blocks) of either a semantic or phonological study task (as in Otten & Rugg, 2001). Each epoch of 83s contained 12 words. By randomly varying the time between words, the correlation between the epoch-related (or state-related) and event-related (or item-related) regressors was reduced to an extent that each effect could be estimated with reasonable efficiency (Chawla et al., 1999). The item-related effects replicated those of Otten and Rugg (2001). More interestingly, a few regions showed a state-related effect that varied linearly with the number of words remembered within each epoch (even though differential item effects had been removed). For the semantic task, these regions were in left ventral prefrontal cortex (-48 +18 +3) and precuneus (-3 -60 +30); for the phonological task, there was one region in dorsal precuneus (0 -54 +60). These state-related encoding effects may reflect, for example, variations across epochs in strategies or levels of sustained attention. Only the left ventral prefrontal region also showed an item-related subsequent memory effect. Interestingly, the state-related activity in this region correlated negatively with subsequent memory, unlike the positive correlation of its item-related activity. One possibility is that a trade-off occurred between the cognitive processes underlying the two effects. These data clearly force further theorising about the nature of different processes that engender good memory.

A final important question is whether the same regions that predict subsequent recollection also reflect that recollection at test (as might be expected from a neural “transfer appropriate processing” account, Blaxton et al., 1996). In the only two studies to compare imaging data acquired at study and test, the regions predicting subsequent recollection and the regions showing recollection-dependent responses did not overlap (Cansino et al., 2002; Henson et al., 1999b). This lack of overlap suggests that the cortical regions predicting subsequent recollection may have more to do with the type of “working memory” operations that lead to better memory (e.g. organisation, Wagner et al., 1999), rather than the type of contextual information associated with those memories (contrary to the “componential” view of episodic memories outlined above). However, further studies are necessary to examine the question of study-test interactions in more detail.

## VII. Conclusion

The most robust repetition effects in recognition memory tests are found in cortical regions, including prefrontal cortex and lateral and medial parietal cortices. Activation of some of these, particularly left anterior prefrontal, left lateral parietal and posterior cingulate, appear to respond most vigorously in association with recollection (at least for words). These activations may reflect re-instatement of study context in working memory. Other prefrontal activations (particularly in dorsolateral regions) are more likely to relate to post-retrieval monitoring of retrieved information (in order, for example, to make a confidence judgment). In other words, these prefrontal activations may reflect decision processes required by recognition memory tasks, rather than memory per se. This is particularly apparent in the study of Herron et al. (in prep), in which some prefrontal old-new effects can become new-old effects simply by varying the proportion of old items.

Activation of MTL regions, as would be expected for neuropsychological studies of amnesia, are less robust. Nonetheless, one interesting pattern beginning to emerge is the association of familiarity with anterior MTL cortex (particularly perirhinal cortex), and association of recollection with the hippocampal formation, as predicted by the model of Aggleton and Brown (1998). An example of this pattern arises in the single dissociation of the Cansino et al. (2002) study, though further studies are required to explicitly test this post hoc observation. One elaboration of this dual-process view is that perirhinal cortex codes the “recency” of perceiving a stimulus (conforming to the memory strength continuum of signal detection models), whereas hippocampus associates that stimulus with the spatiotemporal context in which it was perceived.

Other imaging studies have examined encoding into memory, by investigating the brain regions that predict subsequent recognition memory (particularly subsequent recollection). These regions may depend on the type of processes performed at study, consistent with a proceduralist account of memory. However, some studies have also identified MTL regions that predict subsequent memory, consistent with the structuralist view of a specialised “memory system”. An appealing hypothesis is that episodic memories consist of a number of components that relate to the particular stimulus attributes elaborated at study (depending on the study task), and hence engage different cortical regions according to those attributes (e.g. the semantic or phonological properties of words). However, a necessary condition for these components to “bound” together in a single episodic memory may be that they are associated with an “index” represented in the MTL.

One important avenue for future research will be to relate findings using direct memory tasks with the findings in analogous paradigms using indirect memory tasks (described in the Henson, this volume). In particular, there is a need to dissociate the contributions of implicit and explicit memory to both types of test. For example, as discussed in Section V, the perirhinal new-old effects could represent a form of implicit memory (i.e. the repetition suppression often associated with priming). One way to test this is to perform a recognition memory test that produces sufficient numbers of recognition misses. If the perirhinal cortex response covaries with the participant’s memory judgment (i.e. is deactivated for hits, and possibly false alarms, but not for misses), then that response is more likely to represent a familiarity signal used in recognition memory (Rugg et al., 1998b). If however its response to hits and misses is similar (and different from that to correct rejections and possibly false alarms) – in other words, the response covaries with the objective fact of whether or not the item was repeated – then that

response is more likely to represent an implicit form of memory. Alternatively, one could examine these responses as a function of various manipulations known to affect explicit or implicit memory differently (e.g, changes in study-test modality), or in patients with impaired recognition memory or impaired priming (see Conclusion Section of Henson, this volume). Now that studies of repetition priming and of recognition memory have produced a number of robust findings, future studies might benefit from more careful comparison of repetition effects across these two types of memory test. In any case, the extant data promise an exciting future.

- Aggleton, J. P., & Brown, M. W. (1998). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *The Behavioral and Brain Sciences*.
- Blaxton, T. A., Bookheimer, S. Y., Zeffiro, T. A., Figlozzi, C. M., Gaillard, W. D., & Theodore, W. H. (1996). Functional mapping of human memory using PET: comparisons of conceptual and perceptual tasks. *Canadian Journal of Experimental Psychology*, *50*, 42-56.
- Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). Making memories: brain activity that predicts how well visual experience will be remembered. *Science*, *281*, 185-187.
- Brown, M. W., & Xiang, J. Z. (1998). Recognition memory: neuronal substrates if the judgement of prior occurrence. *Progress in Neurobiology*, *55*, 149-189.
- Buckner, R. L., & Wheeler, M. E. (2001). The cognitive neuroscience of remembering. *Nature Neuroscience Reviews*, *2*, 624-634.
- Cansino, S., Maquet, P., Dolan, R. J., & Rugg, M. D. (2002). Brain activity underlying encoding and retrieval of source memory. *Cereb Cortex*, *12*, 1048-56.
- Chawla, D., Rees, G., & Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nat Neurosci*, *2*, 671-6.
- Desgranges, B., Baron, J.-C., & Eustache, F. (1998). The functional neuroanatomy of episodic memory: the role of the frontal lobes, the hippocampal formation, and other areas. *Neuroimage*, *8*, 198-213.
- Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, *24*, 523-533.
- Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: a selective role for the hippocampus during retrieval. *Nat Neurosci*, *3*, 1149-52.
- Fitzgerald, P. G., & Picton, T. W. (1981). Temporal and sequential probability in evoked potential studies. *Psychophysiology*, *18*, 141-141.
- Fletcher, P. C., & Henson, R. N. (2001). Frontal lobes and human memory: insights from functional neuroimaging. *Brain*, *124*, 849-81.
- Fried, I., Cameron, K. A., Yashar, S., Fong, R., & Morrow, J. W. (2002). Inhibitory and excitatory responses of single neurons in the human medial temporal lobe during recognition of faces and objects. *Cereb Cortex*, *12*, 575-84.
- Gardiner, J. M., Gawlik, B., & Richardson-Klavehn, A. (1994). Maintenance rehearsal affects knowing not remembering: elaborative rehearsal affects remembering not knowing. *Psychonomic Bulletin & Review*, *1*, 107-110.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Hamann, S. B., & Squire, L. R. (1997). Intact perceptual memory in the absence of conscious memory. *Behav Neurosci*, *111*, 850-4.
- Henson, R. N. A. (this volume). Implicit Memory.
- Henson, R. N., Shallice, T., & Dolan, R. J. (1999a). Right prefrontal cortex and episodic memory retrieval: a functional MRI test of the monitoring hypothesis. *Brain*, *122*, 1367-1381.
- Henson, R. N. A., Cansino, S., Herron, J. E., Robb, W. G. K., & Rugg, M. D. (in press). A Familiarity Signal in Human Anterior Medial Temporal Cortex? *Hippocampus*.
- Henson, R. N. A., Rugg, M. D., Shallice, T., & Dolan, R. J. (2000). Confidence in recognition memory for words: dissociating right prefrontal roles in episodic retrieval. *Journal of*

- Cognitive Neuroscience*, 12, 913-923.
- Henson, R. N. A., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. (1999b). Recollection and familiarity in recognition memory: an event-related fMRI study. *Journal of Neuroscience*, 19, 3962-3972.
- Herron, J., Henson, R. N. A., & Rugg, M. D. (in prep).
- Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: measuring recollection. *Journal of Experimental Psychology: General*, 122, 139-154.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Review*, 114, 3-28.
- Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 699-710.
- Kolers, P. A., & Roediger, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, 23, 425-449.
- Mandler, G. (1980). Recognizing: the judgement of previous occurrence. *Psychological Review*, 87, 252-271.
- Maratos, E. J., Dolan, R. J., Morris, J. S., Henson, R. N., & Rugg, M. D. (2001). Neural activity associated with episodic memory for emotional context. *Neuropsychologia*, 39, 910-20.
- Meunier, M., Bachevalier, J., Mishkin, M., & Murray, E. A. (1993). Effects on visual recognition of combined and separate ablations of the entorhinal and perirhinal cortex in rhesus monkeys. *J Neurosci*, 13, 5418-32.
- Otten, L. J., Henson, R. N., & Rugg, M. D. (2001). Depth of processing effects on neural correlates of memory encoding: relationship between findings from across- and within-task comparisons. *Brain*, 124, 399-412.
- Otten, L. J., Henson, R. N., & Rugg, M. D. (2002). State-related and item-related neural correlates of successful memory encoding. *Nat Neurosci*, 28, 28.
- Otten, L. J., & Rugg, M. D. (2001). Task-dependency of the neural correlates of episodic encoding as measured by fMRI. *Cereb Cortex*, 11, 1150-60.
- Rugg, M. D. (2002). Functional imaging of memory. In *Handbook of Memory Disorders (2<sup>nd</sup> Edition)*, Baddeley, A.D., Kopleman, M.D. & Wilson, B.A. (Eds.). John Wiley & Sons, Chichester, England. pp. 57-80.
- Rugg, M. D., Fletcher, P. C., Allan, K., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J. (1998a). Neural correlates of memory retrieval during recognition memory and cued recall. *Neuroimage*.
- Rugg, M. D., Fletcher, P. C., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J. (1996). Differential activation of the prefrontal cortex in successful and unsuccessful memory retrieval. *Brain*, 119, 2073-2083.
- Rugg, M. D., Henson, R. N., & Robb, W. G. (2003). Neural correlates of retrieval processing in the prefrontal cortex during recognition and exclusion tasks. *Neuropsychologia*, 41, 40-52.
- Rugg, M. D., & Henson, R. N. A. (in press). Episodic memory retrieval: an (event-related) functional neuroimaging perspective. In A. E. Parker, E. L. Wilding, & T. Bussey (Eds.), *The cognitive neuroscience of memory encoding and retrieval* : Psychology Press.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998b).

Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*, 595-8.

Rugg, M. D., & Wilding, E. L. (2000). Retrieval processing and episodic memory. *Trends Cogn Sci*, *4*, 108-115.

Schacter, D. L., & Tulving, E. (1994). Memory Systems 1994. In D. L. Schacter & E. Tulving (Eds.) (pp. 1-38). Cambridge, MA: MIT.

Schott, B., Richardson-Klavehn, A., Heinze, H. J., & Duzel, E. (2002). Perceptual priming versus explicit memory: dissociable neural correlates at encoding. *J Cogn Neurosci*, *14*, 578-92.

Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J., & Dolan, R. J. (1994). Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature*, *368*, 633-635.

Strange, B. A., Otten, L. J., Josephs, O., Rugg, M. D., & Dolan, R. J. (2002). Dissociable human perirhinal, hippocampal, and parahippocampal roles during verbal encoding. *J Neurosci*, *22*, 523-8.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, *26*, 1-12.

Wagner, A. D., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998a). Prefrontal cortex and recognition memory: functional-MRI evidence for context-dependent retrieval processes. *Brain*, *121*, 1985-2002.

Wagner, A. D., Koutstaal, W., & Schacter, D. L. (1999). When encoding yields remembering: insights from event-related neuroimaging. *Philos Trans R Soc Lond B Biol Sci*, *354*, 1307-24.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B. R., & Buckner, R. L. (1998b). Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, *281*, 188-191.

Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and Cognition*, *5*, 418-441.

## Figure Legends

### Figure 1.

Schematic of signal detection theory, together with classification of recognition decisions (boxed labels).

### Figure 2.

Data from Henson et al. (1999b) showing regions with greater event-related responses to correct Remember (R) than correct Know (K) decisions to Old words (top row), and the opposite pattern of greater event-related responses to K than R decisions (bottom row). Data are rendered onto the lateral surfaces of a canonical brain. N = correct rejections of New words. Zero level is average across all three response categories.

### Figure 3.

Data from Henson et al. (2000) showing regions with greater event-related responses to hits (Old) than correct rejections (New), collapsed across confidence rating (top row), and greater event-related responses to correct Low than High confidence responses, collapsed across Old/New (bottom row). HN = High confidence New response; LN = Low confidence New response; LO = Low confidence Old response; HO = High confidence Old response. Zero level is average across all three response categories.

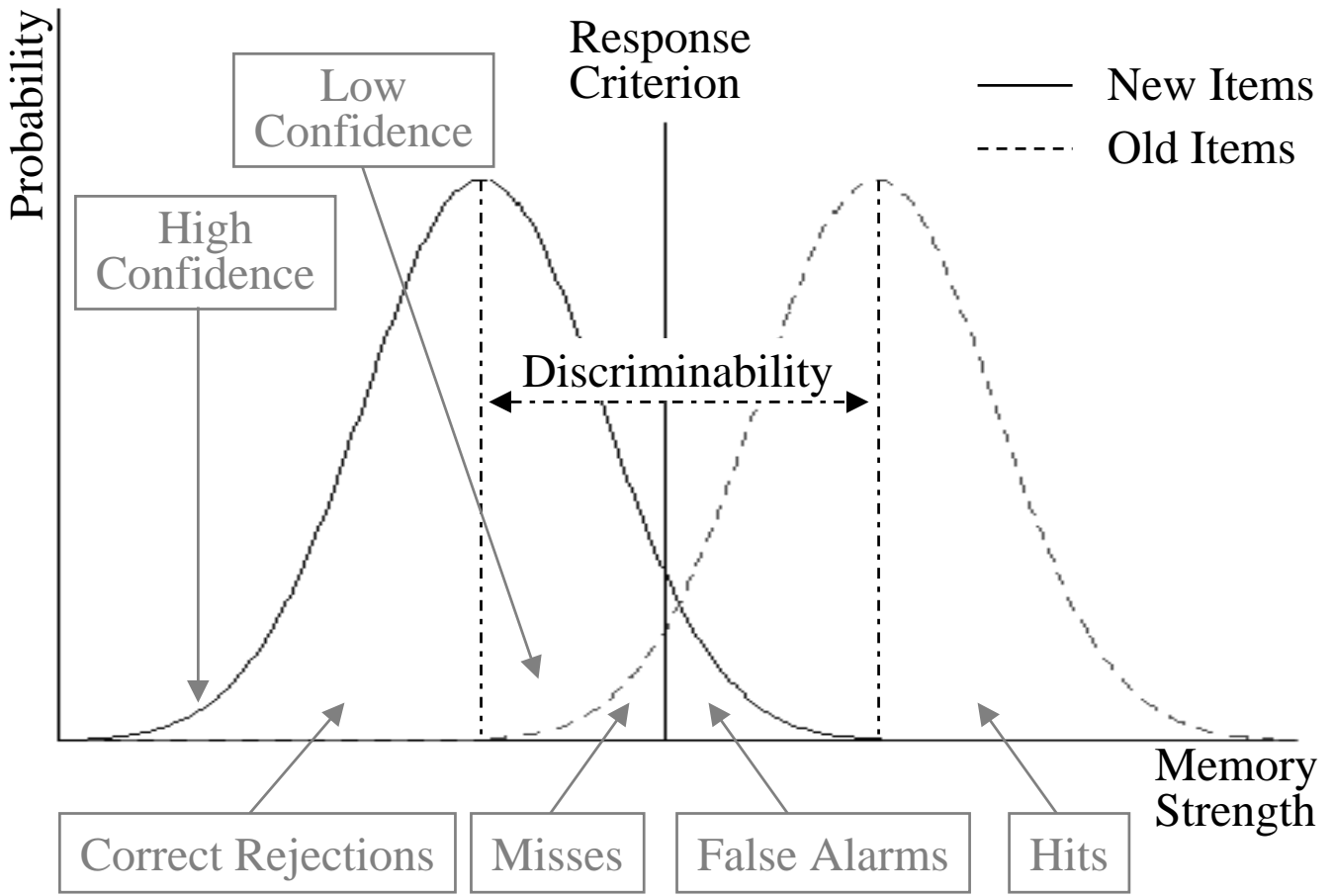
### Figure 4.

Coronal sections through a canonical brain from Cansino et al. (2002), showing regions in anterior medial temporal cortex (left) and hippocampus (right). Below each region is the best-fitting canonical response for correct rejections of new items (red), old items given the correct source (blue) and old items given the incorrect source (green).

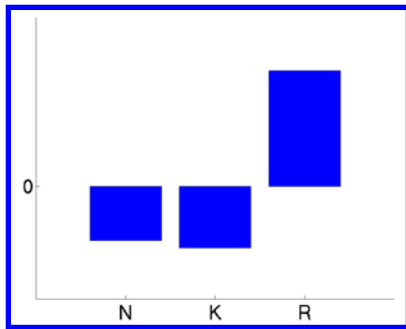
### Figure 5.

Regions from Otten and Rugg (2001) showing differential subsequent memory effects during a semantic task and during a phonological task. Bars show difference between canonical response parameter estimates for subsequently remembered versus subsequently forgotten items for each region under the semantic task (blue) and phonological task (green).

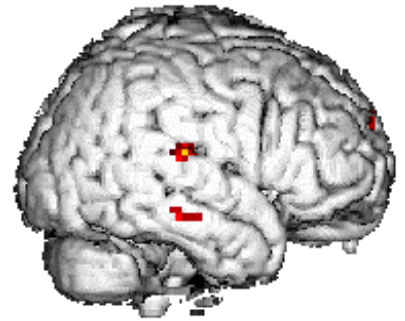
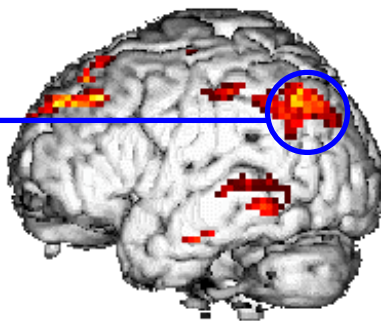




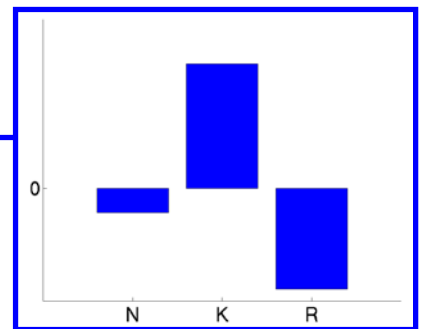
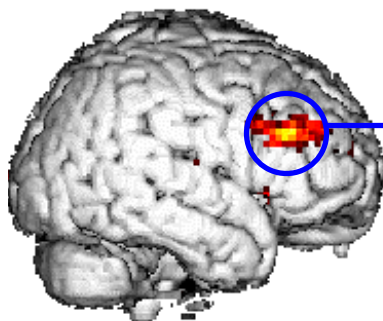
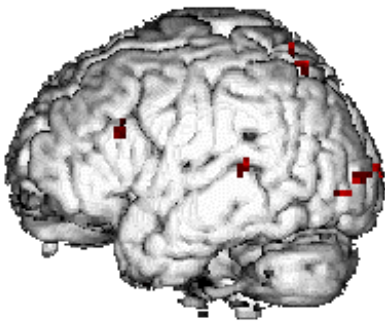
## REMEMBER - KNOW



Left Lateral Parietal

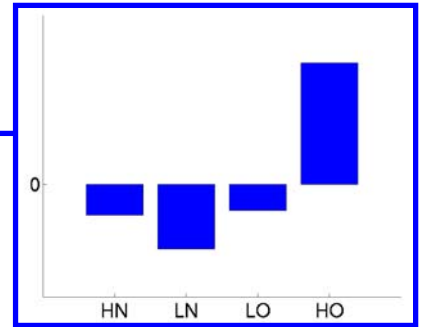
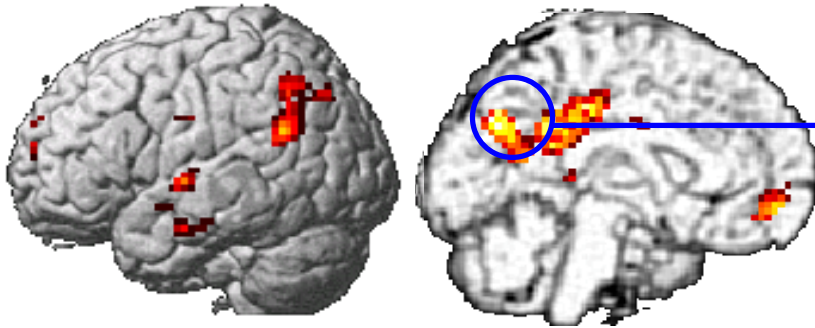


## KNOW - REMEMBER



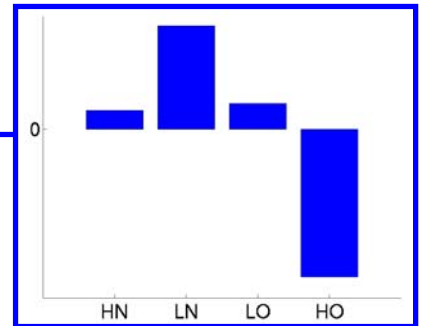
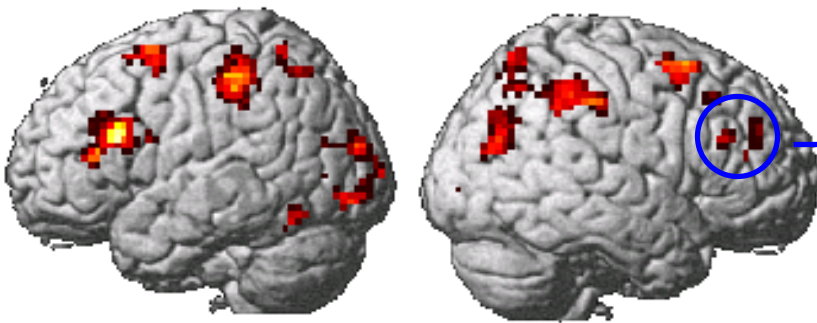
Right Dorsal Prefrontal

# OLD - NEW

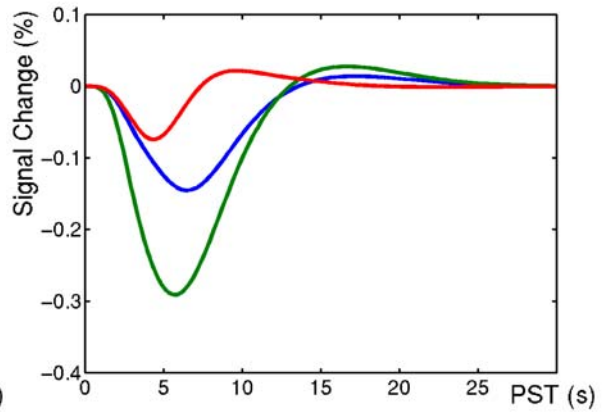
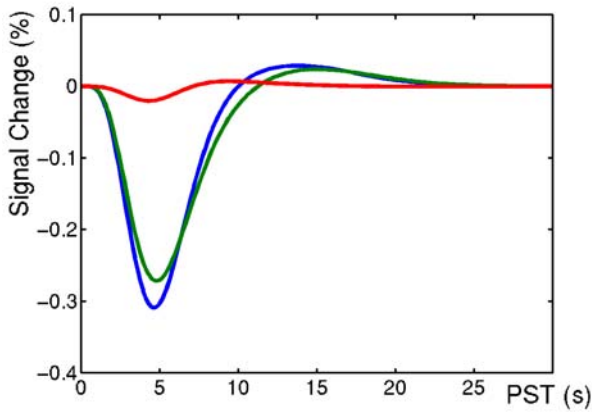
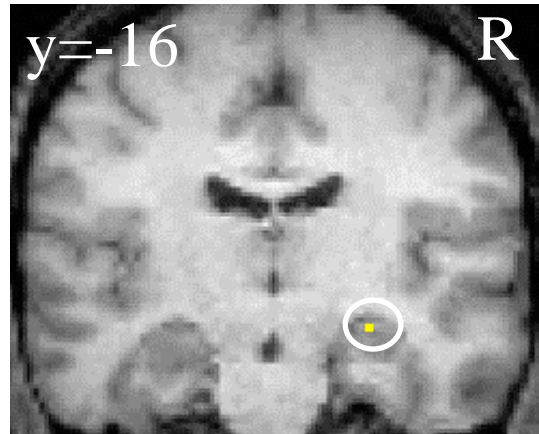
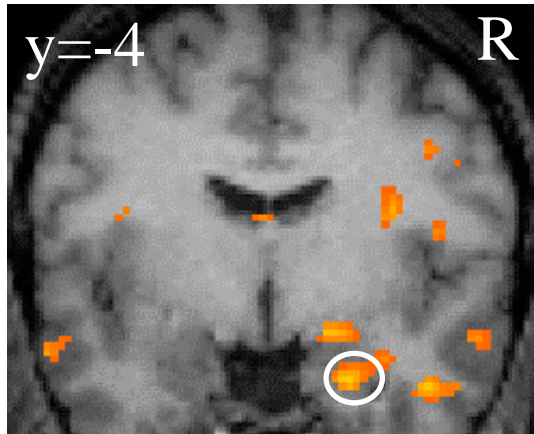


Medial Parietal

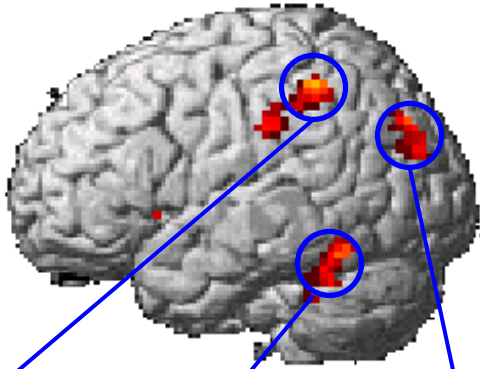
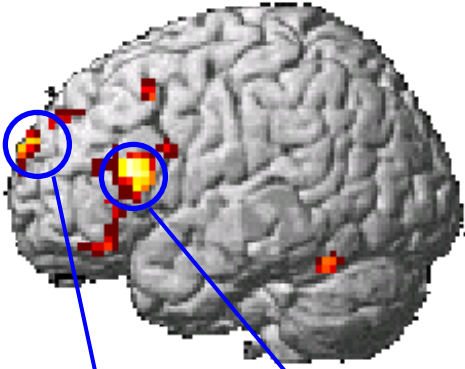
# LOW - HIGH CONFIDENCE



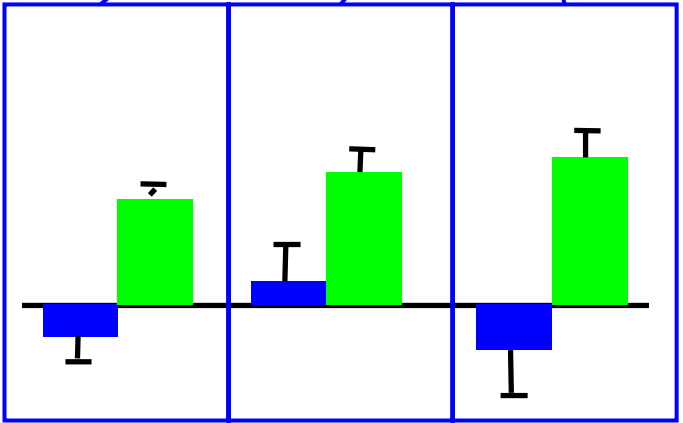
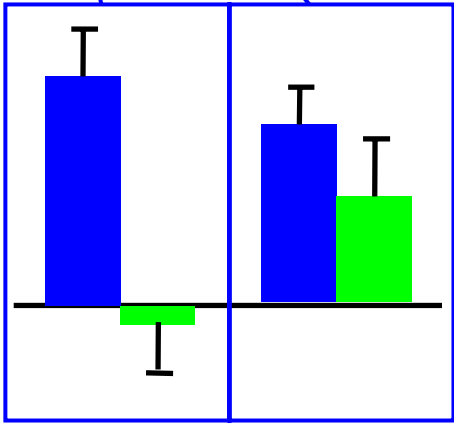
Right Dorsal Prefrontal



**— cor. rej.**      **— cor. source**      **— inc. source**



Subs. Mem. Effect



■ Semantic

■ Phonological