nature neuroscience

A sensory signature that distinguishes true from false memories

Scott D Slotnick & Daniel L Schacter

Human behavioral studies show that there is greater sensory/perceptual detail associated with true memories than false memories. We therefore hypothesized that true recognition of abstract shapes would elicit greater visual cortical activation than would false recognition. During functional magnetic resonance imaging (fMRI), participants studied exemplar shapes and later made recognition memory decisions ("old" or "new") concerning studied exemplars (old shapes), nonstudied lures (related shapes) and new shapes. Within visual processing regions, direct contrasts between true recognition ("old" response to an old shape; old-hit) and false recognition ("old" response to a related shape; related-false alarm) revealed preferential true recognition–related activity in early visual processing regions (Brodmann area (BA)17, BA18). By comparison, both true and false recognition were associated with activity in early and late (BA19, BA37) visual processing regions, the late regions potentially supporting "old" responses, independent of accuracy. Further analyses suggested that the differential early visual processing activity reflected repetition priming, a type of implicit memory. Thus, the sensory signature that distinguishes true from false recognition may not be accessible to conscious awareness.

Memory is often accurate, but everyday experience and laboratory research indicate that it is also prone to various kinds of distortions, errors and illusions[1,2]. Early in the 20th century, F.C. Bartlett showed that when people read stories, they sometimes falsely remembered specific details that were never presented[3]. Much subsequent research has provided evidence for the occurrence of a related form of memory distortion known as false recognition, where people incorrectly claim that they recently encountered a novel object or event[4]. For example, Roediger and McDermott[5], building on earlier work by Deese[6], presented participants with associated word lists (*e.g.*, web, insect, bug, fly, …) that were each related to a critical related but nonstudied lure (*e.g.*, spider). During a subsequent test phase, participants falsely recognized a large proportion of critical lures and expressed high confidence in these inaccurate memories. Similar kinds of robust false recognition effects have been reported for visual items[7]: lists of exemplar shapes were presented at study, and on a later recognition test, participants incorrectly claimed that they previously saw nonstudied related shapes. Despite the high false alarm rates in such experiments, true recognition of previously studied items has been associated with reports of greater contextual and sensory detail than false recognition[8–10].

A fundamental yet poorly understood issue concerns the neural basis of false recognition in relation to true recognition. In an early positron emission tomography (PET) study[11], participants heard semantically associated words at study, and on a subsequent test they made old-new recognition decisions to old words (studied items), related words (critical nonstudied lures) and new words. True recognition–related activity was greater than false recognition–related activity in the left temporoparietal region, which may have reflected memory of auditory/phonological information. However, subsequent studies using similar procedures suggest that such true-false differences depend on the format of the recognition test[9,12].

In a more recent fMRI study[13], two videotaped speakers read lists of semantically associated words and categorized lists to participants during the study phase. During the test phase, participants were scanned as they made old-new recognition memory decisions when presented with old words, related words and new words. They showed robust levels of true and false recognition, both of which were associated with activity in the prefrontal cortex, the parietal cortex and the hippocampus, regions that have been associated with veridical old-new recognition memory in other studies[14–17]. In addition, true recognition, as compared with false recognition, was associated with greater activity in the parahippocampal gyrus, a region that has been associated with contextual processing[18]. Such parahippocampal activity may have reflected true recognition-related contextual memory (for the videotaped speakers).

Based upon the differential activity found during true as compared to false recognition in the previous two studies[11,13], coupled with findings of greater memory for sensory details during true versus false recognition in behavioral studies[8–10], we posited that true recognition is associated with greater sensory reactivation than false recognition. Recent studies examining memory retrieval have provided converging evidence for true recognition-related sensory reactivation of the same cortical regions involved in processing stimulus materials during encoding, including reactivation of motor processing regions during memory for motor sequences[19], reactivation of

Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, Massachusetts 02138, USA. Correspondence should be addressed to S.D.S. (slotnick@wjh.harvard.edu).

Published online 23 May 2004; doi:10.1038/nn1252

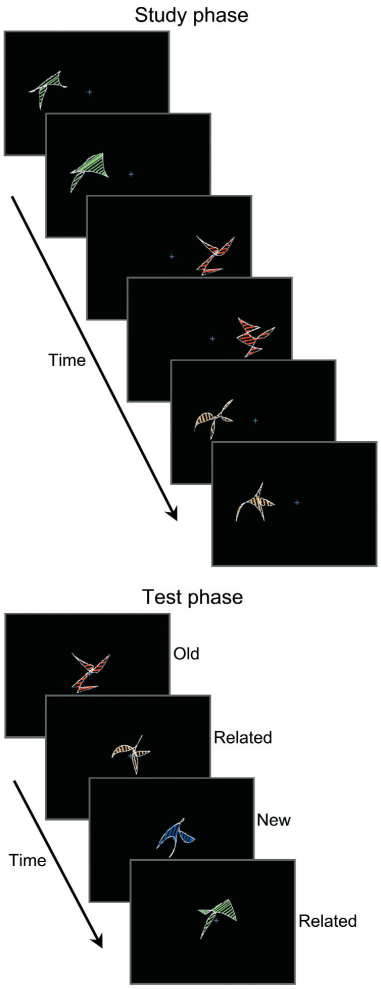

Figure 1 Depiction of experimental protocol (see Methods for details). During the study phase, participants fixated the central cross while sets of exemplar shapes (9 per set, 2 shown) alternated in presentation between the left and right of fixation. Participants were instructed to remember each shape and its spatial location. During the test phase, three types of shapes were presented centrally: (i) old shapes from the study phase, (ii) related shapes that were similar to shapes from the study phase and (iii) new shapes. Participants indicated whether each shape was "old" or "new". True recognition was defined as an "old" response to an old item (old-hit), whereas false recognition was defined as an "old" response to a related item (related-false alarm).

auditory processing regions during memory for sounds[20,21] and reactivation of visual processing regions during memory for pictorial stimuli[17,21–23]. Capitalizing on these findings, we tested our hypothesis in the visual system, given its well-known hierarchical functional-anatomic cortical processing architecture.

In the present event-related fMRI study, we used abstract shapes in an old-new recognition memory task (**Fig. 1**; see Methods). According to our sensory reactivation hypothesis, we expected to observe greater true as compared to false recognition-related visual cortical activity. Furthermore, we attempted to characterize whether sensory reactivation effects were related to conscious recognition of studied materials. Previous behavioral studies of true versus false recognition suggest that sensory memory associated with true recognition reflects greater conscious recognition of sensory/perceptual details. However, other research has shown that retrieval of previously studied perceptual information can be expressed as nonconscious or implicit memory, in the form of repetition priming effects[24]. We attempted to determine whether activity in visual processing regions was specifically related to the subjective sensation that an item had been seen before (conscious recognition), or whether it occurred for all items, regardless of whether they were judged as old or new (*i.e.*, implicit memory[25]). A whole-brain analysis was also conducted to better elucidate the differential and common regions associated with true and false recognition.

Here we report evidence of a functional-anatomic dichotomy between forms of access to late and early visual processing regions: late visual processing regions supported conscious recognition (and were associated with both true and false recognition), whereas early visual processing regions supported implicit memory (and were preferentially associated with true recognition, as compared to false recognition). These results provide direct evidence that previously reported memory-related reactivation in late visual processing regions[17,21–23] is accessible to conscious recognition, which previously has only been assumed. Furthermore, the present results suggest that the previously reported true-greater-than-false activity assumed to reflect sensory or contextual memory[11,13] is largely inaccessible to conscious recognition.

RESULTS

Behavioral results

Participants were able to remember shapes from the study phase: the old-hit rate (p("old"/old) = 63.7 ± 2.9%, mean ± s.e.m.) was significantly greater than the new-false alarm rate (p("old"/new) = 26.2 ± 2.7%; t = 13.1, $P < 0.001$, paired t-test). Critically, the related shapes elicited a robust false recognition effect, as illustrated by a significantly greater related-false alarm rate (p("old"/related) = 55.6 ± 3.0%) than new-false alarm rate ($t = 11.2$, $P < 0.001$, paired t-test). In addition, the old-hit rate and related-false alarm rate were also significantly different ($t = 4.8$, $P < 0.001$, paired t-test), providing some evidence that old and related shapes could be distinguished from one another.

There were no significant reaction-time differences in any of the event-related comparisons, ruling out a 'time on task' explanation of the fMRI results. Specifically, neither the old-hit reaction time (2,223 ± 75 ms) nor the related-false alarm reaction time (2,227 ± 81 ms) was significantly different than the new-correct rejection reaction time (2,155 ± 79 ms; t = 1.18 and t = 1.02, respectively, paired t-test), and the old-hit reaction time was not significantly different than the old-miss reaction time (2,325 ± 123 ms; t = 1.14, paired t-test).

Confirming that there was adequate power to conduct each statistical contrast of interest, there were a sufficient number of observations within all critical response bins for each participant. This included old-hits (range = 50–72, mean = 59.8 ± 2.8), old-misses (range = 23–51, mean = 34.1 ± 2.8), related-false alarms (range = 29–63, mean = 51.7 ± 2.8) and new-correct rejections (range = 46–79, mean = 68.2 ± 3.0).

Distinct neural regions

There were many regions differentially associated with true and false recognition (**Table 1** and **Fig. 2**), including regions previously associated with veridical old-new recognition memory[14–17]. Bearing most directly on our sensory reactivation hypothesis, early visual processing regions (BA17, BA18) showed greater activity during true recognition, as compared to false recognition. Representative event-related activity timecourses from the left fusiform gyrus (BA18; **Fig. 2**, lower middle panel) illustrate both true and false recognition–related increases in activity, but with relatively greater magnitude for true recognition.

Table 1 Neural regions differentially associated with true recognition (old-hits) and false recognition (related-false alarms)

Region	BA	(*x*, *y*, *z*)
True recognition > False recognition		
Left middle frontal gyrus	10	–35, 52, 6
Right middle frontal gyrus	46	46, 33, 23
Right superior frontal gyrus	6	10, 10, 65
Left middle frontal gyrus	6	–29, 2, 59
Right middle frontal gyrus	6	29, 8, 60
Right precentral gyrus	6	22, –11, 67
Right precentral gyrus	4	22, –27, 71
Left caudate	–	–4, 10, 3
Right caudate	–	6, 14, 1
Left superior parietal lobule	7	–16, –53, 58
Right superior parietal lobule	7	14, –66, 58
Left precuneus	7	–10, –69, 42
Left supramarginal gyrus	40	–55, –33, 48
Left fusiform gyrus	18	–19, –85, –13
Left lingual gyrus	18	–16, –88, –13
Right lingual gyrus	18	9, –75, 7
Right cuneus	18	9, –78, 14
Right striate	17	11, –76, 11
False recognition > True recognition		
Right superior frontal gyrus	10	11, 64, 17
Left inferior frontal gyrus	45	–54, 19, 7
Left anterior cingulate gyrus	32	–5, 30, –8
Right anterior cingulate gyrus	24	3, 25, 24
Left insula	13	–37, 6, 8
Right insula	13	38, 13, 10
Left precuneus	7	–4, –54, 33
Left supramarginal gyrus	40	–55, –60, 30
Left superior temporal gyrus	39	–53, –62, 25
Left superior temporal gyrus	22	–60, –23, 4
Left middle temporal gyrus	21	–59, –23, –9

BA, Brodmann area. Talairach coordinates (*x*, *y*, *z*) refer to the center of activation within each region.

In addition to these critical regions, preferential true recognition–related activity was observed in prefrontal regions (BA10, BA46), parietal regions (BA7, BA40) and motor processing regions[26]. The parietal finding replicated a previously reported effect of greater activity for true than for false recognition[13]. Relatively greater false recognition–related activity was also seen in prefrontal regions (BA10, BA45; see event-related timecourses in **Fig. 2**, upper middle panel), parietal regions (BA7, BA40) and the anterior cingulate (BA24, BA32). In addition, false recognition was associated with insular cortex activity (BA13) and left temporal cortex activity (BA21, BA22, BA39), the latter finding being consistent with a previously reported temporal cortex false > true recognition effect[13].

A few regions (BA10, BA7, BA40) were associated with both the true > false recognition contrast and the opposite contrast, a seemingly paradoxical result. However, specific subregions within the prefrontal cortex have been shown to be preferentially associated with certain aspects of encoding and retrieval[27–29], which may also apply to true and false recognition memory. To test this possibility, our BA10 spatial coordinates were compared to those from previous studies of veridical recognition memory (left BA10 range, $x = -21$ to -45, $y = 48$ to 63, $z = 6$ to 21; right BA10 range, $x = 30$ to 40, $y = 46$ to 60, $z = -3$ to 12)[14,15,30–34]. The coordinates of the present true > false recognition–related activity in left BA10 (–35, 52, 6) was well within the range of previous true recognition–related coordinates, whereas the present false > true recognition–related activity in right BA10 (11, 64, 17) was 2.0 cm from the nearest range of previously reported true recognition–related coordinates. The commonly active regions in parietal cortex were also distinct—in left precuneus (BA7), the true recognition coordinates (–10, –69, 42) were 1.8 cm from the false recognition coordinates (–4, –54, 33), supporting the view that relatively more posterior medial parietal cortex is associated with veridical memory[35], and in left BA40, the true recognition coordinates (–55, –33, 48) were 3.2 cm from the false recognition coordinates (–55, –60, 30). These results are consistent with the possibility that distinct sub-regions within BA10, BA7 and BA40, may be preferentially associated with true and false recognition.

Regions associated with conscious recognition

Our key findings of greater activation during true as compared to false recognition in early visual regions BA17 and BA18 are broadly consistent with previous results reporting visual reactivation during true recognition[17,21–23]. Nonetheless, they also differ in at least one respect—previously reported effects occurred in late visual processing regions, most consistently in BA19 and BA37 (within the ventral or 'object' visual processing stream). Importantly, however, we did find both true and false recognition effects within these late visual processing regions (**Table 2**; these and other common regions of

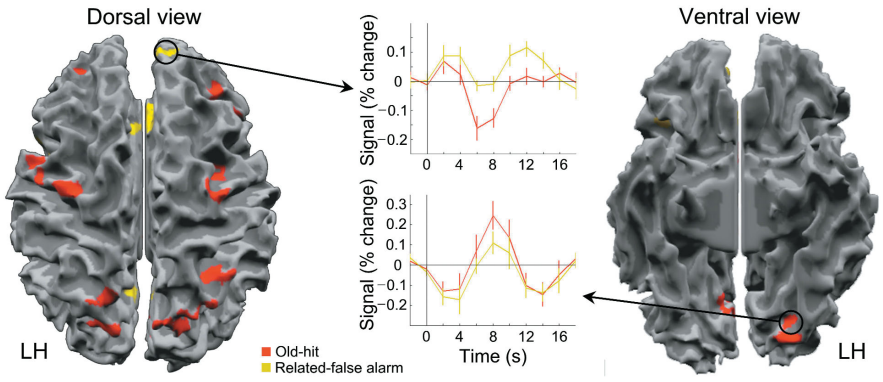

Figure 2 Neural regions differentially associated with true recognition (old-hits > related-false alarms) and false recognition (related-false alarms > old-hits). Activity was projected onto the cortical surface reconstruction of one participant; gyri and sulci are colored light and dark gray, respectively. To the left, the true recognition > false recognition contrast was associated with activity (shown in red) in prefrontal (anteriorly) and parietal cortex (posteriorly), whereas the false recognition > true recognition contrast was associated with less extensive activity (shown in yellow) in prefrontal and parietal cortex with additional reliance on temporal cortex (BA21, BA22, BA39; for a complete list, see **Table 1**). A false recognition > true recognition–related activity time course from anterior prefrontal cortex (BA10) is shown in the upper middle panel (see color key to lower middle; error bars represent s.e.m.). To the right, only the true recognition > false recognition contrast was associated with activity in visual processing regions (BA17, BA18). Lower middle panel, the event-related activity time courses from posterior left fusiform cortex (BA18).

Table 2 Neural regions associated with both true recognition (old-hits > new-correct rejections) and false recognition (related-false alarms > new-correct rejections)

Region	BA	(*x*, *y*, *z*)
Left inferior frontal gyrus	10	–44, 47, –3
Right middle frontal gyrus	10	42, 53, 7
Left inferior frontal gyrus	45	–50, 20, 4
Right middle frontal gyrus	46	54, 32, 21
Left frontal operculum	47	–35, 21, –9
Right frontal operculum	47	35, 23, –6
Left middle frontal gyrus	9	–46, 16, 37
Right middle frontal gyrus	9	52, 14, 34
Left middle frontal gyrus	8	–46, 16, 40
Right middle frontal gyrus	8	27, 21, 45
Left medial frontal gyrus	8	–5, 22, 45
Left middle frontal gyrus	6	–30, 6, 54
Right middle frontal gyrus	6	31, 14, 52
Left anterior cingulate gyrus	32	–7, 38, 21
Left posterior cingulate gyrus	30	–11, –62, 14
Left caudate	–	–8, 6, 5
Right caudate	–	9, 6, 0
Right thalamus	–	5, –17, 12
Left superior parietal lobule	7	–32, –55, 57
Right superior parietal lobule	7	39, –57, 52
Left precuneus	7	–8, –71, 35
Right precuneus	7	6, –67, 48
Left angular gyrus	39	–41, –74, 32
Right angular gyrus	39	43, –63, 40
Left supramarginal gyrus	40	–49, –39, 46
Right supramarginal gyrus	40	42, –49, 45
Left fusiform gyrus	37	–49, –49, –16
Left superior occipital gyrus	19	–43, –76, 28
Left lingual gyrus	18	–8, –87, –11
Right lingual gyrus	18	6, –77, 1
Left striate	17	–5, –88, –1

BA, Brodmann area. Talairach coordinates (*x*, *y*, *z*) refer to the center of activation within each region.

activity are discussed below). Such observations indicate that only early visual regions are more active during true as compared to false recognition, whereas late visual regions are similarly active during true and false recognition. It is possible that late visual processing regions (BA19, BA37) contribute to the conscious experience of remembering, thus supporting "old" responses during both true and false recognition. This possibility is consistent with a recent finding that late visual processing regions are associated with "remember" responses (memory with specific details) but not "know" responses (memory without specific details) during old-new recognition[23]. Indeed, if participants did have conscious access to the differential activity associated with true and false recognition in early visual areas (BA17, BA18), the sensory signature associated with true recognition should have allowed them to selectively endorse old items and correctly reject related items. The high false-recognition rate in the present study shows this was not the case, suggesting participants did not have conscious access to this differential activity in early visual areas.

To directly test the role of these visual processing regions during conscious recognition, we compared activation associated with "old" responses to studied shapes (old-hits) and "new" responses to studied shapes (old-misses), restricting the analysis to the visual processing regions under scrutiny (BA17, BA18, BA19, BA37). The contrast between old-hits and old-misses was employed to reveal those regions associated with conscious recognition, as activity preferentially associated with old-hits was expected to track retrieval content[17,23] (with item type being held constant). By contrast, we expected regions associated with both old-hits and old-misses (with new-correct rejections serving as baseline) to reflect activity independent of conscious awareness–implicit memory[25].

A clear functional dichotomy was observed across visual processing regions (**Table 3**, left column and **Fig. 3a**). While the old-hits > old-misses contrast was associated only with activity in late visual processing regions (BA19, BA37), both old-hits and old-misses were associated with activity in earlier visual processing regions (BA17, BA18). To assess the reliability and generalizability of these results, we conducted a follow-up experiment with 12 additional participants. The protocol was generally similar to the main experiment, with two major differences: single shapes (rather than sets of shapes) were presented during the study phase, and there were no related shapes presented during the test phase (see Methods). Participants were able to remember the shapes, as illustrated by a greater old-hit rate (70.2 ± 2.6%) than new-false alarm rate (38.9 ± 3.3%; $t = 10.0$, $P < 0.001$, paired *t*-test) and there were no significant reaction time differences between old-hits (2,378 ± 150 ms) and old-misses (2,400 ± 144; $t < 1$, paired *t*-test), or between either of these trial types as compared to new-correct rejections (2,354 ± 144; both *t*-values < 1, paired *t*-tests).

Table 3 Regions within BA 17, 18, 19, and 37 associated with old items categorized by response accuracy (old-hits and old-misses)

Main experiment			Follow-up experiment		
Region	**BA**	**(*x*, *y*, *z*)**	**Region**	**BA**	**(*x*, *y*, *z*)**
Old-hits > Old-misses			**Old-hits > Old-misses**		
Left fusiform gyrus	37	–42, –63, –12	Left fusiform gyrus	37	–47, –62, –14
Right fusiform gyrus	37	47, –60, –11	Left fusiform gyrus	19	–42, –76, –12
Left inferior temporal gyrus	37	–58, –53, –9	Left middle occipital gyrus	19	–55, –64, –6
Right inferior temporal gyrus	37	50, –53, –7			
Left middle occipital gyrus	19	–45, –71, –7			
Right middle occipital gyrus	19	50, –68, –5			
Old-hits & Old-misses			**Old-hits & Old-misses**		
Left lingual gyrus	18	–8, –88, –12	Left middle occipital gyrus	18	–36, –86, 3
Left cuneus	18	–4, –87, 13	Right lingual gyrus	18	14, –79, –3
Right cuneus	18	3, –80, 16	Left striate	17	–10, –86, 6
Left striate	17	–5, –89, 1	Right striate	17	20, –79, 2
Right striate	17	16, –82, 5			

BA, Brodmann area. Talairach coordinates (*x*, *y*, *z*) refer to the center of activation within each region.

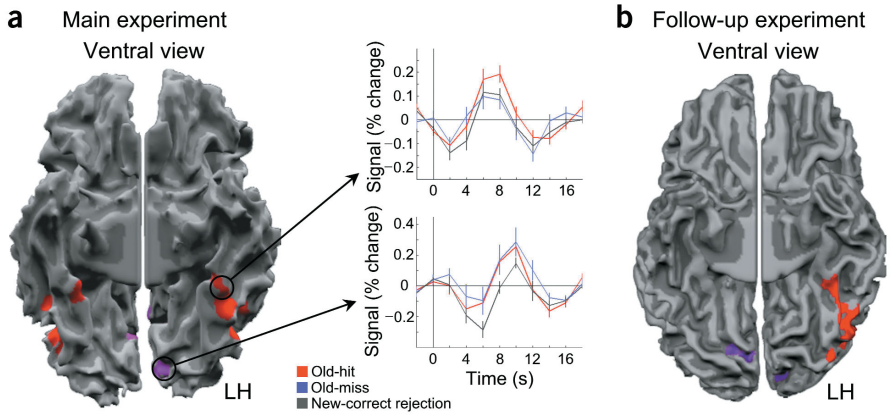

Figure 3 Ventral brain regions commonly (magenta) and differentially (red) associated with old-hits and old-misses. (**a**) Main experiment. (**b**) Follow-up experiment. In both experiments, the old-hits > old-misses contrast was associated with more anterior (late) visual processing regions (BA19, BA37), whereas the regions associated with both old-hits (old-hits > new-correct rejections) and old-misses (old-misses > new-correct rejections) were associated with posterior (early) visual processing regions (BA17, BA18). Event-related activity time courses from the left fusiform gyrus (BA37) and left cuneus (BA18) of the main experiment are shown in the upper and lower middle panels, respectively (see color key to lower middle).

The protocol was specifically designed to maximize the number of critical trials that were included in the analysis in an effort to increase experimental power. Indeed, compared with the initial experiment, there were relatively more observations of old-hits (range = 100–161, mean = 129.8 ± 5.0), old-misses (range = 27–90, mean = 55.8 ± 5.2) and new-correct rejections (range = 26–72, mean = 57.1 ± 3.7). Motivated by this increase in power, we also used a more stringent statistical threshold in the old-hits > old-misses contrast to minimize type I error ($P < 0.01$, corrected for multiple comparisons to $P < 0.001$; see Methods). The follow-up experiment produced the identical pattern of results observed in the main experiment, where late visual processing regions (BA19, BA37) were associated with the old-hits > old-misses contrast, while early visual processing regions (BA17, BA18) were associated with both old-hits and old-misses (**Table 3**, right column and **Fig. 3b**).

The results of both the main experiment and follow-up experiment provide direct evidence that conscious recognition may be specifically associated with activity within late visual processing regions. This interpretation is consistent with our findings of similar true and false recognition–related activity in late visual areas (BA19, BA37; see below), thus supporting "old" responses when participants were presented with either old or related items. It is also consistent with our findings of differential true and false recognition–related activity in early visual areas (BA17, BA18), to which there may be little or no conscious access, such that this activity failed to provide a basis for correctly rejecting a majority of related shapes.

Common neural regions

In addition to the late visual processing regions noted above, several other regions were similarly active for both true and false recognition (**Table 2** and **Fig. 4**). The common pattern of prefrontal (BA10, BA45, BA46, BA47, BA8, BA9) and parietal cortex activity (BA7, BA39, BA40) was similar to that previously associated with veridical old-new recognition[14–17], and replicates a previous study implicating these regions during false recognition[13]. Representative event-related activity timecourses were extracted from a parietal region (BA 39/40; Talairach coordinates –39, –55, 36; **Fig. 4**, upper middle panel) that has been associated with "old" responses[17]. The anterior cingulate (BA32), posterior cingulate (BA30), motor processing regions and thalamus have also been associated with veridical old-new recognition[15,26,31,36].

Although we also expected activity in the hippocampus to be associated with both true and false recognition[13], significant hippocampal activity was associated only with true recognition (**Fig. 5**, left; Talairach coordinates –18, –27, –9). However, the fact that the true > false recognition contrast was not associated with activity in the hippocampus suggested there was false recognition–related activity in this region, but at a sub-threshold level. Indeed, extraction of event-related activity timecourses from the true recognition-related hippocampal region-of-interest (**Fig. 5**, left) showed activity increases associated with both true and false recognition (**Fig. 5**, right). A post-hoc analysis on these activity timecourses revealed significant true and false recognition-related increases 6 s following stimulus onset (old-hit vs. new-correct rejection, $t = 2.20$, $P < 0.05$; related-false alarm vs. new-correct rejection $t = 1.86$, $P < 0.05$; old-hit vs. related-false alarm, $t < 1$). Thus, the hippocampus was, in fact, associated to some extent with both true and false recognition in the present study, consistent with previous results[13].

DISCUSSION

In the present study, a large number of neural regions were associated with both true and false recognition, confirming previously reported

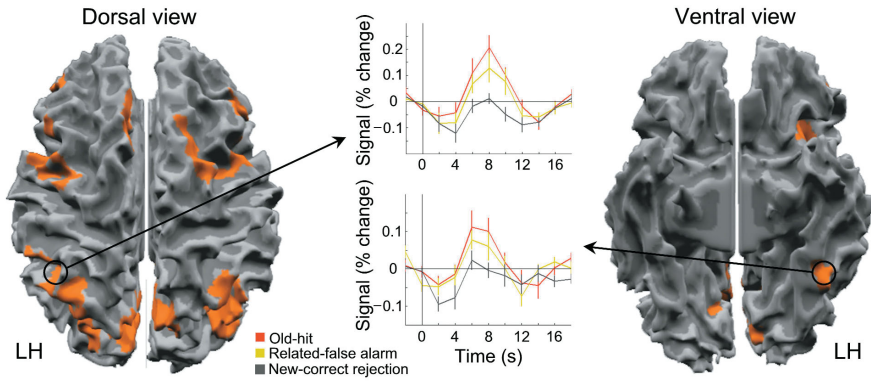

Figure 4 Neural regions associated with both true recognition (old-hits > new-correct rejections) and false recognition (related-false alarms > new-correct rejections). To the left, the common regions of activity (orange) within frontal and parietal cortex can be observed (for a complete list, see **Table 2**). Event-related time courses were extracted from parietal cortex near BA 39/40, as shown in the upper middle panel (see color key to lower middle). To the right, a number of visual processing regions (BA17, BA18, BA19, BA37) were active during both true and false recognition. Lower middle panel, event-related activity time courses from the left fusiform gyrus (BA37).

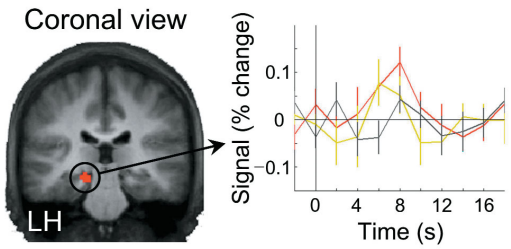

Figure 5 Hippocampal activity associated with true and false recognition. Left, activity within the hippocampus associated with true recognition (old-hits > new-correct rejections) is shown in red. Right, event-related time courses from this region-of-interest show increases in activity associated with both true and false recognition (see **Fig. 4** for color key).

activations in the prefrontal cortex, parietal cortex and hippocampus[13], including a left parietal region recently implicated in endorsing items as "old"[17]. Additionally, true recognition and false recognition were similarly active in late visual processing regions (BA19, BA37), also in line with previous results[17,21–23]. This common neural architecture associated with both true and false recognition may reflect the neural basis of responding "old" to old studied and related nonstudied items. There were also a number of neural regions differentially associated with true and false recognition, although participants were not able to consistently capitalize upon such differential activity to inform their behavioral response (or they would have correctly rejected most of the related items). Of note, false > true recognition–related activity was greater in the left temporal cortex, including Wernicke's area (BA22); this language processing–related activity[37] may represent a previously reported bias toward verbal retrieval strategies associated with false memory, as compared to true memory[8]. Of greatest importance, and in support of our sensory reactivation hypothesis, true > false recognition–related activity was observed in early visual processing regions (BA17, BA18).

Consistent with previous behavioral reports of greater sensory and perceptual details during true as compared to false recognition[8–10], the differential activity observed in early visual regions might have been interpreted as a neural correlate of conscious recognition for sensory/perceptual information associated with previously studied shapes. However, subsequent analyses showed that old-hits and old-misses evoked similar levels of activity within BA17 and BA18. In line with previous arguments[25], these findings suggest that activation within these early visual processing regions reflect a nonconscious or implicit form of memory retrieval. Although previous event-related potential (ERP) studies have reported old-miss–related activity over parietal and occipital scalp[25,38] (with new-correct rejections serving as baseline), this is the first time, to our knowledge, that the precise neural correlates of old-miss–related activity in an old-new recognition memory paradigm have been reported. We propose that the old-hit and old-miss-related increases in activity within these regions reflect the operation of a particular type of implicit memory: repetition priming[24]. Previous studies have shown that priming of familiar objects is typically associated with lower activation relative to baseline[39–44], perhaps raising doubts concerning our hypothesis. Importantly, however, old item-related (*i.e.*, priming-related) increases in activity have been reported during implicit memory tasks that have employed unfamiliar faces or novel objects as stimuli[45–47]. The novel object studies[46,47] can be considered particularly suitable as a basis for comparison, because that stimulus set consisted of simple geometric line drawings (possible and impossible shapes) that were of comparable complexity to the shapes employed in the present study. As in the present study, these experiments showed priming related increases in visual processing regions. Thus, it is plausible that the increases we observed with novel shapes reflect a manifestation of repetition priming.

It could be argued that some old-misses occurred when participants were unsure about their prior responses; common activations for old-hits and old-misses in early visual processing regions therefore might have been driven by the occurrence of "unsure" old-miss responses that were close to the "old" response threshold. That is, rather than representing priming/implicit memory as we have argued, the present old-miss–related activity may instead have been based on weak explicit memory for studied shapes. To address this possibility, we examined the old-miss–related activity from the left cuneus (BA18; **Fig. 3**) in the main experiment and the right lingual gyrus (BA18) in the follow-up experiment, binned according to "unsure"/"sure" confidence (only participants with at least ten responses of each item type were included in the analysis; $n = 8$ and $n = 7$ for the main experiment and follow-up experiment, respectively). At 8 s after stimulus onset, the magnitude of old-miss-sure–related activity was similar to the old-miss-unsure–related activity in both experiments (main $t < 1$ and follow-up $t = 1.06$), and greater than the new-correct rejection–related activity, although the latter effect was marginally significant in each experiment alone (main $t = 1.42$ and follow-up $t = 1.67$, both P-values < 0.10), the joint probability of observing two such outcomes is $P < 0.05$ (calculated using the Fisher technique[48]; see Methods). We observed a similar pattern of old-miss–related activity time courses in the other early visual processing regions reported in both experiments, which argues against a threshold account of the old-miss effects.

It should also be noted that previous studies in which participants studied pictures of common objects have not reported early visual cortical reactivation during memory retrieval[17,21–23]. These findings are also consistent with our priming hypothesis, as those studies presented word cues during the test phase. This mismatch between study and test modalities likely precluded the possibility of observing repetition priming, which is highly sensitive to perceptual overlap between study and test stimuli[24]. In contrast, the present design repeated items during the study and test phases, thus allowing for perceptual priming effects.

Although the true and false recognition–related activity differences observed here in early visual areas seem to depend largely on implicit processes, the possibility that some of this information could be consciously accessed cannot be ruled out. Indeed, the relatively small but significant behavioral differences associated with true and false recognition in the present study may have been partly attributable to a limited degree of conscious access to such differential activity within early visual areas. It has been shown that the demand for detailed information during memory retrieval can be manipulated[49,50]; thus, stimulus or task factors that lead to more careful scrutiny of the detailed visual characteristics during memory retrieval may foster greater conscious access to activity within earlier visual processing regions. Such conscious access to the differential visual activity within these regions may effectively serve to reduce the occurrence of false recognition.

METHODS

Main experiment: participants. Fourteen right-handed Harvard undergraduates with normal or corrected-to-normal visual acuity were enrolled in the experiment. The training protocol was approved by the Harvard University Institutional Review Board and the imaging protocol was approved by the Massachusetts General Hospital Institutional Review Board. All participants

provided informed consent for both portions of the experiment. Two participants were excluded from the analysis due to a paucity of responses (4% and 45% of total items) in addition to chance levels of old-new recognition memory performance. The twelve participants (7 females) included in the analysis were 20 ± 0.57 years of age (mean ± s.e.m.).

Task and procedure. Participants first completed a training session at Harvard University identical to the subsequent imaging session at the Massachusetts General Hospital imaging center (except stimuli were never repeated). Both sessions consisted of three study-test phases. Each study list consisted of 144 shapes (16 sets of 9 exemplars; see **Fig. 1** and description of stimulus construction in **Supplementary Methods** online) where exemplar sets alternated in presentation between the left and right visual field. Each shape was presented for 2.5 s with an inter-trial interval (ITI) of 3 s. Participants were instructed to remember each shape and its spatial location, while always maintaining fixation at the central cross. Test lists consisted of 96 shapes (16 sets of 2 studied exemplars/old items, 2 related shapes/related items and 2 nonstudied shapes/new items) presented at the center of the screen for 2.5 s with a random ITI ranging from 4–12 s. For each shape, participants made two judgments using their left hand: (i) whether it was "old and on the left," "old and on the right" or "new" and (ii) whether they were "unsure" or "sure" of the preceding response. Participants were instructed to respond quickly but without sacrificing accuracy, while always maintaining central fixation. There were no differences between spatial source accuracy for old items (percentage correct = 73.9 ± 1.9; 50% representing chance performance) and related items (73.3 ± 1.9% correct; $t < 1$, paired t-test), and no differences between confidence attributions associated with these item types (p("sure"/old) = 48.3 ± 6.3%; p("sure"/related) = 46.2 ± 6.5%; $t < 1$, paired t-test); thus, to maximize power, the present analysis was conducted without regard to source accuracy or confidence at test (unless otherwise noted). As such, behavioral responses were classified as "old" or "new". Old, related and new lists of shapes were counterbalanced across participants using a Latin square design.

Follow-up experiment. The follow-up experiment was nearly identical to the main experiment, such that subsequent details are limited to those aspects in which the experiments differed. Twelve participants (7 females, 21 ± 0.75 years of age) that had not taken part in the main experiment completed the follow-up experiment and were included in the analysis. Immediately after one study-test phase for training, six study-test phases were conducted in an imaging session. Each study phase consisted of 32 shapes presented to the left or right of fixation (16 on the left and 16 on the right, with random assignment of spatial location). Each test phase included the 32 studied shapes from the study phase and 16 new shapes; participants made the same responses as in the main experiment.

Image acquisition and data analysis. The imaging protocol and data analysis methodology for the main experiment and follow-up experiment were nearly identical; therefore, what follows will be a single description, with any differences being noted. Images were acquired using a 3-tesla Siemens Allegra MRI scanner. A multiplanar rapidly acquired gradient echo (MP-RAGE) sequence was used to acquire high-resolution anatomic volumes (TR = 30 ms, TE = 3.3 ms, 128 slices, 1×1×1.33 mm resolution) and a T2*-weighted echo planar imaging (EPI) sequence was used for functional imaging (TR = 2,000 ms, TE = 30 ms, 64×64 acquisition matrix, 26–30 slices for the main experiment, 30 slices for the follow-up experiment, 4.5 mm isotropic resolution). The variable number of slices in the main experiment was due to initial technical limitations of the scanner (when only 26 slices could be acquired with a TR of 2,000 ms); nonetheless, whole brain coverage was achieved for all participants. Stimuli were projected onto a screen at the superior end of the scanner bore that was viewed through an angled mirror affixed to the head coil. Under these viewing conditions, shapes were contained within a bounding square 5.5° of visual angle in width, with the closest edge offset 3° of visual angle from fixation during study. All participants completed a single anatomic scan, three study runs and three test runs in the main experiment, and six study-test runs in the follow-up experiment.

Unless otherwise specified, pre-processing and data analysis was conducted using SPM99 software (Wellcome Department of Cognitive Neurology). Images were first slice-time corrected and motion corrected. The variable number of slices in the main experiment required entering participant-specific slice order information during the slice-time correction procedure, but all other aspects of the analysis were identical. After this, using custom software written in Matlab (MathWorks, Inc.), each run was temporally high-pass filtered by removing linear, quadratic, cubic and quartic components, which is approximately the same as passing frequencies greater than 0.0026 Hz, and all runs were concatenated. Images were then transformed into a standard space, including re-sampling at 3 mm isotropic resolution. Images were not spatially smoothed.

Event-related analysis was conducted using a random-effect general linear model approach. For each participant, square-waves representing onsets and durations for each of the event types (*e.g.*, in the main experiment, old-hits, old-misses, related-false alarms, related-correct rejections, new-false alarms and new-correct rejections) were convolved with a canonical hemodynamic response function. For each voxel, these hemodynamic response models were simultaneously fit to the activation time course, thus producing a beta-weight (*i.e.*, model amplitude) associated with each event type. In the second level of analysis, only those voxels with beta-weight differences that were statistically consistent across participants were considered active, as assessed using a one-tailed t-test.

Unless otherwise noted, only individual voxel activity significant to $P < 0.05$ is reported. Furthermore, activity was cluster extent threshold corrected for multiple comparisons to $P < 0.01$, as specified via Monte Carlo simulations[16]. The cluster extent threshold procedure relies on the fact that given spurious activity or noise (voxel-wise type-I error), the probability of observing increasingly large (spatially contiguous) clusters of activity systematically decreases. Thus, a cluster extent threshold can be enforced to ensure an acceptable probability of cluster-wise type-I error. Each simulation consisted of 1,000 independent iterations where the brain volume was modeled by a 64×64×30 matrix (*i.e.*, acquisition matrix by slice; a 64×64×26 matrix produced identical results), and activity at each voxel was modeled by assigning a normally distributed random number (with a mean of zero and a variance of 1). A voxel activity threshold was then applied to achieve the desired individual voxel type-I error rate (p-value). After this, the spatial extent of each cluster was computed (minimum = 1 voxel) and the number of clusters with the same spatial extent tallied. The total number of each cluster extent, across all iterations, was used to compute an overall probability of observing that cluster extent, relative to the total number of clusters. Using these probabilities, a cluster extent threshold was selected such that the total probability of observing all clusters of that extent threshold or larger was equal to the desired cluster-wise type I error rate (i.e. the p-value corrected for multiple comparisons). The simulation associated with the aforementioned individual voxel and corrected p-values resulted in a cluster extent threshold of 31 contiguous resampled voxels (837 mm^3), an extent threshold that was enforced for all contrasts unless otherwise stated. In comparison, the analytically derived cluster extent threshold from SPM99 to correct for multiple comparisons at $P < 0.05$ was 57 voxels, illustrating that this analytical method is relatively strict when compared to the currently used Monte Carlo method.

When the union of two contrasts is displayed (*e.g.,* old-hits > new-correct rejections and related-false alarms > new-correct rejections), the former contrast was computed first to define the region-of-interest (with 31 contiguous voxels) within which the spatial extent of activity associated with the latter contrast (i.e. the joint activity of the two contrasts) was used to compute a corrected p-value. In the present experiment, the joint probability of observing an active voxel was $P = 0.018$, computed using the Fisher's technique[48] where the corresponding chi-square value was calculated using the equation

$$\chi^2 = -2\mathrm{Ln}(p_1 p_2)$$

where p_1 and p_2 both equal 0.05 and there are $2 \times n = 4$ degrees of freedom. This joint p-value defined the individual voxel level of significance that translated to a correction for multiple comparisons to $P < 0.001$ (with 31 contiguous voxels), except for the old-hits & old-misses comparison of the follow-up experiment that was corrected for multiple comparisons to $P < 0.05$ (with 9 contiguous voxels).

Event-related activity time courses were extracted from spherical regions of interest (with a radius of 6 mm) using custom software written in Matlab. Event-related time courses were produced in a similar manner to that used in

ERP studies—for a given event type, the activation time course (% signal change as a function of time) for every trial of that event was first computed, time-locked to stimulus onset, and then this activity was averaged across trials to produce the activation time course reported. Each time course was baseline-corrected to produce an average of 0% signal change from 0 and 2 s before stimulus onset. Linear trend removal was also applied to each activity time course. To determine the expected maximum amplitude of the event-related response, a squarewave representing an event duration of 2.5 s was convolved with a canonical hemodynamic response function (HRF), as a function of time (t), of the form described by

$$\mathrm{HRF} = \left(\frac{t-\delta}{\tau}\right)^2 \exp^{-\left(\frac{t-\delta}{\tau}\right)}$$

where $\delta = 2.5$ and $\tau = 1.25$. As the maximum of the resulting hemodynamic response was observed between 6 and 8 s following stimulus onset, *post-hoc* statistical tests applied to event-related time courses were limited to those two time points. It should also be noted that an initial undershoot is sometimes observed in the event-related time courses (*e.g.*, 2 s after stimulus onset in the upper middle panel of **Fig. 3a**); this is an artifact of the relatively short ITIs used in the present study, where the initial event-related decrease in activity is due to a return to baseline of the previously occurring event (**Supplementary Methods**). As these undershoots are due to activity from a mixture of previous event types, it is reasonable to assume that they will have no differential effects on the event-related time courses. To confirm this, using custom software written in Matlab, we ran a deconvolution analysis with an impulse response model to reconstruct the non-overlapping event-related time courses. As expected, the initial undershoot was absent from the reconstructed event-related time courses. Furthermore, the pattern of results and statistical inferences associated with the reconstructed event-related time courses were identical to those associated with the actual event-related time courses; therefore, only the results of the actual event-related time courses are reported in detail.

For viewing purposes, fMRI results were imported into BrainVoyager (Brain Innovation), transformed into Talairach space, and projected onto a gray/white matter surface reconstruction of one participant. It is important to note that the neuroanatomical configuration of this participant is not the same as the group average; as such, the activations observed on this surface should be taken only as a reflection of the group results (see **Tables 1–3** for precise activation coordinates).

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We thank L. Moo for helpful discussions and use of her BrainVoyager software. This work was supported by grants MH-NS60941 from the National Institute of Mental Health and AG08441 from the National Institute on Aging.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 3 March; accepted 22 March 2004
Published online at http://www.nature.com/natureneuroscience/

1. Schacter, D.L. The seven sins of memory. Insights from psychology and cognitive neuroscience. *Am. Psychol.* **54**, 182–203 (1999).
2. Roediger, H.L. Memory illusions. *J. Mem. Lang.* **35**, 76–100 (1996).
3. Bartlett, F.C. *Remembering: A Study in Experimental and Social Psychology* (Cambridge Univ. Press, London, 1932).
4. Underwood, B.J. False recognition produced by implicit verbal responses. *J. Exp. Psychol.* **70**, 122–129 (1965).
5. Roediger, H.L. & McDermott, K.B. Creating false memories: remembering words not presented in lists. *J. Exp. Psychol. Learn. Mem. Cogn.* **21**, 803–814 (1995).
6. Deese, J. On the prediction of occurrence of particular verbal intrusions in immediate recall. *J. Exp. Psychol.* **58**, 17–22 (1959).
7. Koutstaal, W., Schacter, D.L., Verfaellie, M., Brenner, C. & Jackson, E.M. Perceptually based false recognition of novel objects in amnesia: effects of category size and similarity to category prototypes. *Cogn. Neuropsychol.* **16**, 317–341 (1999).
8. Schooler, J.W., Gerhard, D. & Loftus, E.F. Qualities of the unreal. *J. Exp. Psychol. Learn. Mem. Cogn.* **12**, 171–181 (1986).
9. Mather, M., Henkel, L.A. & Johnson, M.K. Evaluating characteristics of false memories: remember/know judgments and memory characteristics questionnaire compared. *Mem. Cogn.* **25**, 826–837 (1997).
10. Norman, K.A. & Schacter, D.L. False recognition in younger and older adults: exploring the characteristics of illusory memories. *Mem. Cogn.* **25**, 838–848 (1997).
11. Schacter, D.L. *et al.* Neuroanatomical correlates of veridical and illusory recognition memory: evidence from positron emission tomography. *Neuron* **17**, 267–274 (1996).
12. Schacter, D.L., Buckner, R.L., Koutstaal, W., Dale, A.M. & Rosen, B.R. Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. *Neuroimage* **6**, 259–269 (1997).
13. Cabeza, R., Rao, S.M., Wagner, A.D., Mayer, A.R. & Schacter, D.L. Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proc. Natl. Acad. Sci. USA* **98**, 4805–4810 (2001).
14. Buckner, R.L., Koutstaal, W., Schacter, D.L., Wagner, A.D. & Rosen, B.R. Functional-anatomic study of episodic retrieval using fMRI. *Neuroimage* **7**, 151–162 (1998).
15. Lepage, M., Ghaffar, O., Nyberg, L. & Tulving, E. Prefrontal cortex and episodic memory retrieval mode. *Proc. Natl. Acad. Sci. USA* **97**, 506–511 (2000).
16. Slotnick, S.D., Moo, L.R., Segal, J.B. & Hart, J. Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Cogn. Brain Res.* **17**, 75–82 (2003).
17. Wheeler, M.E. & Buckner, R.L. Functional dissociation among components of remembering: control, perceived oldness, and content. *J. Neurosci.* **23**, 3869–3880 (2003).
18. Bar, M. & Aminoff, E. Cortical analysis of visual context. *Neuron* **38**, 347–358 (2003).
19. Nyberg, L. *et al.* Reactivation of motor brain areas during explicit memory for actions. *Neuroimage* **14**, 521–528 (2001).
20. Nyberg, L., Habib, R., McIntosh, A.R. & Tulving, E. Reactivation of encoding-related brain activity during memory retrieval. *Proc. Natl. Acad. Sci. USA* **97**, 11120–11124 (2000).
21. Wheeler, M.E., Petersen, S.E. & Buckner, R.L. Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proc. Natl. Acad. Sci. USA* **97**, 11125–11129 (2000).
22. Vaidya, C.J., Zhao, M., Desmond, J.E. & Gabrieli, J.D.E. Evidence for cortical encoding specificity in episodic memory: memory-induced re-activation of picture processing areas. *Neuropsychologia* **40**, 2136–2143 (2002).
23. Wheeler, M.E. & Buckner, R.L. Functional-anatomic correlates of remembering and knowing. *Neuroimage* **21**, 1337–1349 (2004).
24. Tulving, E. & Schacter D.L. Priming and human memory systems. *Science* **247**, 301–306 (1990).
25. Rugg, M.D. *et al.* Dissociation of the neural correlates of implicit and explicit memory. *Nature* **392**, 595–598 (1998).
26. Henson, R.N.A., Shallice, T., Gorno-Tempini, M.L. & Dolan, R.J. Face repetition effects in implicit and explicit memory tests as measured by fMRI. *Cereb. Cortex* **12**, 178–186 (2002).
27. Tulving, E., Kapur, S., Craik, F.I.M., Moscovitch, M. & Houle, S. Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proc. Natl. Acad. Sci. USA* **91**, 2016–2020 (1994).
28. Buckner, R.L. Beyond HERA: contributions of specific prefrontal brain areas to long-term memory retrieval. *Psychon. Bull. Rev.* **3**, 149–158 (1996).
29. Wagner, A.D., Maril, A., Bjork, R.A. & Schacter, D.L. Prefrontal contributions to executive control: fMRI evidence for functional distinctions within lateral prefrontal cortex. *Neuroimage* **14**, 1337–1347 (2001).
30. Haxby, J.V. *et al.* Face encoding and recognition in the human brain. *Proc. Natl. Acad. Sci. USA* **93**, 922–927 (1996).
31. Henson, R.N.A., Rugg, M.D., Shallice, T. & Dolan, R.J. Confidence in recognition memory for words: dissociating right prefrontal roles in episodic retrieval. *J. Cogn. Neurosci.* **12**, 913–923 (2000).
32. Donaldson, D.I., Petersen, S.E. & Buckner, R.L. Dissociating memory retrieval processes using fMRI: evidence that priming does not support recognition memory. *Neuron* **31**, 1047–1059 (2001).
33. Rugg, M.D., Henson, R.N.A. & Robb, W.G.K. Neural correlates of retrieval processing in the prefrontal cortex during recognition and exclusion tasks. *Neuropsychologia* **41**, 40–52 (2003).
34. Velanova, K. *et al.* Functional-anatomic correlates of sustained and transient processing components engaged during controlled retrieval. *J. Neurosci.* **23**, 8460–8470 (2003).
35. Buckner, R.L., Raichle, M.E., Miezin, F.M. & Petersen, S.E. Functional anatomic studies of memory retrieval for auditory words. *J. Neurosci.* **16**, 6219–6235 (1996).
36. Konishi, S., Wheeler, M.E., Donaldson, D.I. & Buckner, R.L. Neural correlates of episodic retrieval success. *Neuroimage* **12**, 276–286 (2000).
37. Price, C.J. The anatomy of language: contributions from functional neuroimaging. *J. Anat.* **197**, 335–359 (2000).
38. Walla, P., Endl, W., Lindinger, G., Deecke, L. & Lang, W. Implicit memory within a word recognition task: an event-related potential study in human subjects. *Neurosci. Lett.* **269**, 129–132 (1999).
39. Buckner, R.L. *et al.* Functional anatomic studies of explicit and implicit memory retrieval tasks. *J. Neurosci.* **15**, 12–29 (1995).
40. Buckner, R.L. *et al.* Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron* **20**, 285–296 (1998).

41. Squire, L.R. *et al.* Activation of the hippocampus in normal humans: a functional anatomical study of memory. *Proc. Natl. Acad. Sci. USA* **89**, 1837–1841 (1992).
42. Schacter, D.L. & Buckner, R.L. Priming and the brain. *Neuron* **20**, 185–195 (1998).
43. Wiggs, C.L. & Martin, A. Properties and mechanisms of perceptual priming. *Curr. Opin. Neurobiol.* **8**, 227–233 (1998).
44. Dobbins, I.G., Schnyer, D.M., Verfaellie, M. & Schacter, D.L. Cortical activity reductions during repetition priming can result from rapid response learning. *Nature* **428**, 316–319 (2004).
45. Henson, R., Shallice, T. & Dolan, R. Neuroimaging evidence for dissociable forms of repetition priming. *Science* **287**, 1269–1272 (2000).
46. Schacter, D.L. *et al.* Brain regions associated with retrieval of structurally coherent visual information. *Nature* **376**, 587–590 (1995).
47. Uecker, A. *et al.* Neuroanatomical correlates of implicit and explicit memory for structurally possible and impossible visual objects. *Learn. Mem.* **4**, 337–355 (1997).
48. Fisher, R.A. *Statistical Methods for Research Workers*. edn. 14 (Hafner, New York, 1973).
49. Schacter, D.L., Israel, L. & Racine, C. Suppressing false recognition in younger and older adults: the distinctiveness heuristic. *J. Mem. Lang.* **40**, 1–24 (1999).
50. Dodson, C.S. & Schacter, D.L. Aging and strategic retrieval processes: reducing false memories with a distinctiveness heuristic. *Psychol. Aging* **17**, 405–415 (2002).