Published in final edited form as:
*Neuron*. 2011 March 10; 69(5): 1002–1014. doi:10.1016/j.neuron.2011.02.013.

# Uncovering Camouflage: Amygdala Activation Predicts Long-Term Memory of Induced Perceptual Insight

**Rachel Ludmer**[1], **Yadin Dudai**[1], and **Nava Rubin**[2,*]

[1]Department of Neurobiology, The Weizmann Institute of Science, Rehovot, Israel

[2]Center for Neural Science, New-York University

## Abstract

What brain mechanisms underlie learning of new knowledge from single events? We studied encoding in long-term memory of a unique type of one-shot experience, induced perceptual insight. While undergoing an fMRI brain scan, participants viewed degraded images of real-world pictures where the underlying objects were hard to recognize ('camouflage'), followed by brief exposures to the original images ('solution'), which led to induced insight ("Aha!"). A week later, participants' memory was tested; a solution image was classified as 'remembered' if detailed perceptual knowledge was elicited from the camouflage image alone. During encoding, subsequently remembered images enjoyed higher activity in mid-level visual cortex and medial frontal cortex, but most pronouncedly in the amygdala, whose activity could be used to predict which solutions will remain in long-term memory. Our findings extend the known roles of amygdala in memory to include promoting of long-term memory of the sudden reorganization of internal representations.

## Introduction

A substantial part of the knowledge that we acquire in real-life is a consequence of a one-time exposure to an event, yet the brain mechanisms that underlie this type of rapid learning are largely unknown. While the prevalent example of single-event knowledge acquisition is episodic memory (Roediger et al., 2007; Tulving, 1983), another type of real-life single-event learning is insight: the sudden realization of a solution to a problem (Hebb, 1949; Kohler, 1925). Although insight is most often discussed in the context of cognitive tasks such as problem solving (Kaplan and Simon, 1990; Sternberg and Davidson, 1995), abrupt improvements in performance, as well as the subjective 'Aha!' experience characteristic of insight, can also be observed in perception (Porter, 1954; Rubin et al., 1997; Rubin et al., 2002). The sudden realization of the solution may happen spontaneously, but it can also be induced by an external cue, both in cognitive problem solving (Maier, 1931) and in perception. Readers may be able to experience induced perceptual insight for themselves by viewing Figure 1, which was generated by degrading a real-world picture, taking a few moments to try to identify the underlying scene, and then turning to Figure 2 (next page), which shows the original image. Upon re-exposure to the degraded image, or 'camouflage' (Figure 1), many observers report perceiving a compelling depiction of the underlying scene -- just moments after the very same image appeared as a meaningless collection of ink blots.

In daily life, information that results from moments of insight is, almost by definition, incorporated in long-term memory: once we have realized a new way to solve a problem, or to perform a task better and faster, we are not likely to forget that insight easily. But what is

*Corresponding author; Nava.Rubin@nyu.edu.

the neural basis of this long-lasting nature of insight? Other forms of learning typically require long training periods and many repeated trials, as has been observed in sensory and perceptual learning (e.g., Gauthier et al., 1997; Karni and Sagi 1991; Seitz and Watanabe 2009), motor learning (e.g., Newell and Rosenbloom 1981) and rote-learning in animals (e.g., Stevens and Savin, 1962). These time-scales accord well with the long held idea that incorporation of new knowledge into long-term memory involves synaptic modifications that require gradual processes, sometimes over weeks or months (Dudai, 2004; Hebb, 1949; Martin et al., 2000; Squire and Kandel, 1999). Yet recent findings indicate that modifications of schemas may form and stabilize quickly (Tse et al., 2007). The retention of insights in memory may therefore provide another avenue to study neural events that support the rapid formation of long term memories.

Induced perceptual insight offers several attractive characteristics as a laboratory model for learning that results from real-life insightful transitions. It allows the experimenter to induce the representational transition fairly reliably at predictable moments, with the presentation of a hint (the original image) for a brief amount of time -- an advantage of particular value when investigating encoding in the fMRI environment And although the transition to the new, insightful perceptual state was externally induced, rather than occurring spontaneously, it often invokes a similar sense of an ‘Aha!’ moment. How the moment of insight came about is obviously of central importance for studies that are primarily concerned with the mental and/or neural processes that give rise to spontaneous insight (e.g., Bowden et al., 2005; Jung-Beeman et al., 2004; Kounios et al., 2008). However, unlike those previous studies, our aim here was to study the neural correlates of memory retention of insightful solutions. In this context, induced perceptual insight offers another important methodological advantage: it is possible to generate a large set of camouflage images and their associated solutions, and expose observers to such large collections of puzzle-solution pairs within, say, an hour -- thus obtaining multiple induced insight events in a time-frame which lends itself well to fMRI scanning (Dolan et al., 1997).

Many observers feel that the perceptual transition they had just underwent was so dramatic that they are going to remember the solution for a long time thereafter. When presented with a single such exemplar (e.g., the dog in Figure 1), the declarative memory of the distinct encoding event may serve as a cue that facilitates reconstruction of the insightful solution (e.g., when encountering this article again you may remember there was a dog in the camouflage image and, if it does not pop-out, search for it). But what will be the fate of the camouflage solutions in terms of their retention in memory when observers are exposed to many of them (say, 30) in one session? Would they remember all of the solutions? This seems unlikely. On the other hand, it is possible that they would remember the solutions to a good fraction of those images. If so, what determines which solution images are retained in memory, and which are not? In particular, can one identify patterns of brain activity that occur during the realization of a solution that could predict the memory outcome of this solution? This is the question we set out to answer in this study, by employing a subsequent memory paradigm, similar to that used in exploring brain mechanisms of encoding of other types of event memory (Brewer et al., 1998; Hasson et al., 2008; Paller et al., 1987; Wagner et al., 1998).

We began by generating a large set of candidate images by systematically degrading pictures of real-world scenes, and evaluating them in pilot experiments. A degraded image was classified as ‘a good camouflage’ if enough observers were unable to recognize the hidden object prior to exposure to the solution (original, un-degraded) image, and yet endorsed it as a perceptually compelling rendition of the original after exposure to the latter image. Next, we tested new groups of subjects for memory retention of the solutions in behavioral experiments. Subjects participated in two sessions. In the first, *‘study’* session, subjects were

first shown each of a set of camouflage images and given an opportunity to report if they recognized the hidden object. They were then exposed to the solution (original) image, and finally back to the camouflage. Subjects were not asked to remember the solutions nor told that the experiment was related to memory. They returned at a later, pre-arranged time for a second, *'test'* session. They were shown the same set of camouflage images, intermingled with a smaller set of novel images, and asked to identify the hidden object in each of the images in turn. This time, however, they were not shown the solution at any stage. Instead, if they made a correct identification (e.g., "a dog" in Figure 1), subjects were given a follow-up question that probed the detail and vividness of their perception (e.g., "where is the nose of the dog?"). Note that our test procedure therefore probed memory for the *content* of the induced insight event ("what is hidden in this camouflage image?"), not the episodic memory of the event itself ("do you remember seeing the solution?").

We used fMRI scanning to compare between brain activity for camouflage images whose solutions were subsequently retained in memory and the activity for the camouflages whose solutions were forgotten. Since we were interested specifically in activity differences during the moment of induced insight – i.e., during presentation of the solution in the first, *'study'* session – subjects were scanned during this session only. The behavioral tests indicated that participants retained many of the solutions in long-term memory, but also forgot a sizable fraction. Importantly, different participants tended to remember different sub-sets of images, and therefore whatever differential fMRI activity we find could not be attributable to differences in the stimulus sets.

We found that activity in several brain regions was correlated with subsequent long-term memory of the solution. Most prominent was the finding that activity in the amygdala during the moment of insight predicted long-term memory retention of the solution. In fact, we were able to use amygdala activity to predict subsequent memory on a trial-by-trial basis in a new group of participants. The role of the amygdala in emotional memory is well established (e.g., McGaugh, 2004; Phelps and LeDoux, 2005). Note, however, that none of the camouflage images in our experiments were in and of themselves emotional, though the internal evaluation of the insightful transition might have incited emotion. This suggests that the amygdala plays an important role also in encoding events whose importance is not given *a priori* and externally, but rather evaluated internally. We also found activity correlated with subsequent memory in mid-level visual cortex and medial frontal cortex. The Discussion provides an account of how activity in those regions may be related to internal evaluative processes while viewing subsequently remembered solutions, and to the subjective "Aha!" experience that often accompanies insight.

## Results

Three experiments are reported here: each consisted of a *'study'* session in which subjects were exposed to camouflage images and their solutions, and a *'test'* session in which subsequent memory of the camouflage solutions was tested. Figure 3 presents the protocol used in *Experiment 2* (whole-brain fMRI; see Supplemental Material Figs. S1 and S3 for the slightly different protocols used in Experiments 1 and 3.) *Experiment 1* provided behavioral measures of the retention in memory of the camouflage solutions over time (no fMRI scanning). In Experiments 2 and 3 the *'study'* session was performed while new groups of participants were undergoing fMRI scanning: whole-brain in *Experiment 2* and a higher-resolution scan focused on the amygdala in *Experiment 3*. In both Experiments 2 and 3 the *'test'* session was performed outside the magnet, a week after the *'study'* session. .

## Experiment 1: Memory of the acquired solutions to camouflage images

Participants first completed a *'study'* session, exposing them to the camouflage images followed by their solutions. Participants saw 30 images out of the total set of 40 (selected randomly for each participant). For each image, the participants indicated whether they perceived the camouflaged object spontaneously prior to the solution. Spontaneously recognized images were excluded from the memory analysis (this resulted in the exclusion of different images for different participants; see below). To assess memory retention over time of those images that were not recognized spontaneously, four different groups of participants were administered a *'test'* session after four different time lags: 15 minutes, 1 day, 1 week and 3 weeks. Participants were again presented with the 30 images from the *'study'*, intermixed with 10 novel images that they were not previously exposed to. Perception of the object embedded in the camouflage was tested first by a 'Multiple-Choice' question, and if a correct answer was given, also by a requirement to indicate the location of a specific feature in the scene ('Grid' task; Fig. S1). Figure 4 presents the memory performance of the different groups: of the images that were not spontaneously recognized during *'study'*, (88±3)% were correctly identified in the multiple choice test 15 min after the completion of the *'study'*, declining to (76±3)% in the 1 day lag group, (57±4)% in the 1 week and (58±4)% in the 3 weeks group (top panel). The correct response in the Grid task, which more faithfully ascertains the vividness of recognition, was (66±5)%, (46±4)%, (36±5)% and (33±5)% after 15 min, 1 day, 1 week and 3 weeks, respectively (bottom panel). There was no significant difference in performance between the 1 week group and the 3 weeks group. Thus, if the solution to a camouflage image is retained a week after seeing it, it is retained to essentially to the same degree also three weeks afterwards.

Might the performance during the *'test'* reflect a learning set or skill acquisition of the task, rather than stimulus-specific memory of the camouflage images and their associated solutions? This can be addressed by examining performance on the 10 camouflage images not seen during the *'study'* session. In all four time-lag groups, performance was significantly better on images that were presented in the *'study'* vs. novel images. This differential performance cannot be attributed to differences in the images attributes, since each participant saw a different subset of 30 camouflage images during *'study'*, drawn randomly from the total of 40 images. Moreover, no significant difference was found between the performance of the different time groups on the novel images (Figure 4, open symbols; Kruskal-Wallis ANOVA by ranks), indicating that the degradation in performance over time on the images seen at *'study'* was not due to a general decline in task performance.

The spontaneous recognition rate in the *'study'* session was (34±3)%. There was no significant difference in spontaneous recognition between the four different time-groups. This level is similar to the multiple choice correct recognition of novel images during *'test'* (Figure 4) and a dependent samples t-test showed no significant difference between the performance in the two tasks, suggesting that there was no general learning of the task above and beyond the stimulus-specific learning.

Importantly, there was no subset of images which accounted for the majority of the remembered images. We calculated the frequency distribution of the Grid task correct responses per image, and the resulted distribution did not significantly differ from the normal distribution (Shapiro-Wilk $p=0.07$). To test if there was an effect of image content on subsequent memory, we performed a Kruskal-Wallis ANOVA on correct recognition per image, grouping images by their content (a human figure, an insect, an animal, an object, a face, or a complex scene). There was no effect of content on subsequent memory performance.

## Experiment 2: Brain correlates of the encoding of subsequent memory of the acquired solution

On the basis of the results of *Experiment 1*, that showed similar memory performance after 1 week and 3 weeks, we decided to test subsequent memory 1 week after performing fMRI scanning during the *'study'* session in *Experiment 2* (see Figure 3 for the protocol and the notation of its stages).

### Behavioral Performance

**Performance in Experiment 2 was similar to that of Experiment 1:** The participants in *Experiment 2*, who performed the *'study'* session in the scanner, recognized spontaneously (27±3)% of the camouflages. In the *'test'* session 1 week later, they provided a correct response to (56±4)% of the camouflages in the Multiple-Choice test and (44±5)% in the Grid task. (Here as elsewhere, spontaneously recognized images were excluded in calculating memory performance.) There was no significant difference between the memory performance of the participants in *Experiment 2* and those tested 1 week after *'study'* in *Experiment 1*. In addition, spontaneous recognition was reproducible across the *'study'* and *'test'* sessions: For images reported as spontaneously recognized during the *'study'*, the correct response in the *'test'* was (85±4)% in the Multiple-Choice test, and (78±6)% in the Grid task. Importantly, and as in *Experiment 1*, there was no subset of images which accounted for the majority of the remembered trials across participants, or significant content effects. These results attain special importance for the fMRI analysis, since any difference in BOLD activity that we may find during *'study'* between images that were subsequently remembered and those that were not remembered would not be attributable to content differences in the images themselves.

**False alarms did not affect subsequent memory:** For some images, participants had false-alarms: they pressed the button to indicate identification of the hidden object during the first presentation of the camouflage image (*CAM1*, Figure 3A), but after seeing the solution (*SOL*), they indicated that they did not actually identify the underlying object correctly (QUERY stage, cf. Figure 3A). False-alarms constituted 23% of the camouflage trials that participants indicated as non-identified in *QUERY*. The group performance in the *'test'* Grid task for false-alarm images (i.e. correct identification of the solution in the *'test'* despite having a false alarm during the *'study' CAM1*) was 44%, the same as the mean performance for all non-identified images, showing no apparent effect of false-alarms on subsequent memory. Those images were therefore included in the subsequent memory analyses.

**fMRI Results—**Our aim was to uncover brain regions in which activity during the *'study'* was correlated with subsequent acquired recognition of the object embedded in the camouflage image. Hence the *'study'* trials were classified based on the behavioral performance as follows: Trials in which the camouflage was reported as spontaneously identified (i.e., when the participant pressed 'Yes' at the *QUERY* stage during '*study*') were labeled *SPONT*. The remaining trials were classified based on performance during the *'test'* session: Those for which the solution was remembered a week later were labeled REM, and those for which the solution was not remembered were labeled *NotREM*. Only images that were answered correctly at both the Multiple-Choice and the Grid task at *'test'* were labeled REM in the subsequent memory analysis. (See Experimental Procedures for further analyses made to validate this choice.)

We report results from a select set of regions of interest (ROIs), as well as a whole-brain GLM analysis. In the GLM analysis, each stage in the trial (*CAM1, SOL, CAM2*) was considered as a separate condition, resulting in 9 conditions: *CAM1-REM, CAM1-NotREM, CAM1-SPONT, SOL-REM* etc. Similarly, in the ROI analyses presented below time-course

data from each of the stages were treated separately according to the behavioral performance.

**Activity in the left amygdala correlated with subsequent memory—**The amygdala was selected as an ROI in an analysis aimed to delineate the regions that were mostly engaged during the presentation of the camouflage solution (i.e., during the period of induced perceptual insight), by contrasting *SOL* vs. baseline activity for all trials, regardless of recognition and/or memory outcome of the trial. (See *Experimental Procedures* subsection *Regions of interest Experiment 2*.) In addition to the amygdala, this contrast also revealed extensive activations in visual and frontal cortices (Figure 5A; for the full list of activations see Supplemental Material, Table S1; visual ROIs were defined using an independent localizer data, see below).

Figure 5B presents the event-triggered average time course activity in the amygdala ROI during the first presentation of the camouflage image (*CAM1*, left panel) and the solution (*SOL*, right panel). During *SOL* the left amygdala showed a significantly higher activation for subsequently remembered images (*REM)* than for the images that were not remembered a week later during *'test'* (*NotREM*). In the right amygdala, activation for *REM* images was also higher than for *NotREM* ones, however the difference was not significant (see Supplemental Material Figure S3). We did not observe significant subsequent memory effects or identification effects in the amygdala during *CAM1* or *CAM2*.

**Activity in visual cortex correlated with subsequent memory—**Four visual cortical ROIs were delineated using data from the 'object localizer' functional scans (contrasting responses to pictures of everyday objects with scrambled versions of the same objects; see *Experimental Procedures*). Two were sub-regions of the lateral occipital complex, the LO (in and around the lateral occipital sulcus) and the pFs (posterior fusiform sulcus), and the others were the CoS (collateral sulcus) and the EarlyVis (in and around the calcarine fissure) ROIs. (See Supplemental Material Figure S2 and Table S1 for anatomical loci.)

We hypothesized that the LOC would show higher activity (i) for *SPONT* events in comparison to trials in which the camouflage was not identified, during the *CAM1* phase of the trials (presentation of the camouflage prior to the solution); (ii) for *REM* events, compared with *NotREM* events, during the *SOL* phase of the trials (presentation of the camouflage alternating with the solution). The first hypothesis is straightforward given the extensive evidence that the LOC plays a key role in human object recognition (Malach et al., 1995, Grill-Spector et al., 2000). The second hypothesis was based on the idea that subsequent memory is more likely in trials when the underlying object is perceived more vividly (after exposure to the solution). This should be observable as higher LOC activity in those trials, compared with trials when the camouflage image was perceived by the participant as giving only a poor portrayal of the solution image. Analysis of the event-triggered average time-course data from our (independently acquired and delineated) LOC ROIs confirmed both hypotheses and, at the same time, exposed interesting differential effects within its sub-regions (Figure 6). Specifically, during *SOL* the right LO (top panels), but not the pFs (bottom panels), showed significantly higher activity for trials that were subsequently remembered compared with trials whose solution was not remembered.

**Activity in medial frontal and parietal cortices correlated with subsequent memory—**A multi-subject voxel-by-voxel subsequent memory contrast was conducted, comparing *SOL-REM* with *SOL-NotREM* trials. This unveiled in addition to clusters of voxels in the lateral-occipital area, foci of subsequent memory correlated activation during *SOL* mainly in left medial prefrontal regions (mPFC, BA 9 and BA10), in the anterior

cingulate cortex (ACC) and in the precuneus (Figure 7; the full list of significant foci of activation is provided in Supplemental Material, Table S1).

**No correlation with subsequent memory was detected in the hippocampus—** Since the hippocampal formation is commonly implicated in multiple types of memory tasks, and also since we found activation in the hippocampus when contrasting *SOL* trials vs. baseline trials (albeit in a small cluster of voxels, see Supplemental Material, Table S1), we delineated hippocampus ROIs (head, body and tail) based on anatomical landmarks. Although the hippocampus ROIs do show some BOLD response in *SOL*, we did not find subsequent memory differential activation in any of these hippocampal ROIs.

## Experiment 3: Using amygdala BOLD data to predict subsequent memory of the acquired solution

On the basis of the results of *Experiment 2*, we ran a third experiment aimed at using fMRI data from a *'study'* session to predict memory performance at a '*test*' to be done one week later. The protocol was slightly different than that of *Experiment 2* (in the *'study' CAM1* was 6 sec instead of 10 sec, and *CAM2* was removed, see *Experimental Procedures* and Supplemental Material Figure S3; the *'test'* session was identical).

**Behavioral Performance—**The participants in *Experiment 3*, who performed the *'study'* session in the scanner, and saw 40 images (instead of 30 as in *Experiments 1 & 2)*, recognized spontaneously (34±8)% of the camouflages. In the *'test'* session 1 week later, they provided a correct response to (42±15)% of the camouflages in the Multiple-Choice test and (27±15)% in the Grid task. Again, and as in *Experiments 1 & 2*, images that participants reported they recognized spontaneously were not included in the analysis. There was no significant difference between the memory performance in the Grid task of the participants in *Experiment 3* and those tested 1 week after the *'study'* in *Experiment 1* (two tailed *t*-test, independent samples; p= 0.24), though there was a difference in the performance in the Multiple-Choice test (p= 0.025), which might be due to the larger image set used in the *'study'*. There was no significant difference in the spontaneous recognition during *'study'*.

**fMRI Analysis—**In *Experiment 2* BOLD activity in the left amygdala correlated pronouncedly with subsequent long term memory performance. Thus in *Experiment 3* we set out to test if amygdala activity during the *'study'* may serve to predict subsequent recognition of camouflage images at the *'test'* on a trial-by-trial basis. Hence we scanned the participants while they performed the *'study'* using a high resolution EPI, resulting in 2*2*2mm voxels, keeping the same TR (2 sec). The scan did not cover the whole brain, but had our region of interest – the amygdala – in the center of the FOV (see Supplemental Material Figure S3). Trials were first classified based only on the *'study'* session behavior as follows: Trials in which the camouflage was reported as spontaneously identified (i.e., when the participant pressed 'Yes' at the *QUERY* stage), were labeled *SPONT*. The rest of the trials in which the camouflage was reported as not identified spontaneously, were labeled *NotIdentified*. We then used the *SOL* vs. baseline contrast, as was done in *Experiment 2*, to delineate the subject-specific amygdala ROIs which we *a priori* set out to test. Subsequent memory information was not used at this stage to avoid circularity when choosing the voxels whose data is used for prediction. Next, we calculated the area under the curve for the peak time points of each *NotIdentified* trial. The trials were sorted by this measure, and following the results of the previous experiments, the top 40% of the sorted trials list were predicted to be subsequently remembered, while the rest were predicted not to be remembered.

**Left amygdala activity predicted subsequent long-term memory of the camouflage solution:** When we compared the above described prediction with the actual performance of

the participants at the *'test'* the average hit rate of the prediction (i.e. the number of trials in which the image was predicted to be remembered, and was indeed recognized at the Grid task one week later, as a fraction of the total number of *REM* trials) was (0.548±0.127). The average false-alarm rate of the prediction (i.e. the number of trials in which the image was predicted to be remembered yet was not recognized at the Grid task one week later, as a fraction of the total number of *NotREM* trials) was (0.312±0.052). The average d-prime for the prediction was (0.628±0.445). The hit rate vs. false-alarm rate relation per subject is depicted in Figure 8. As in *Experiment 2* the right amygdala also showed higher activity in *REM* trials than in *NotREM* ones. Yet again that difference was much smaller than in the left amygdala. The average hit rate, false-alarm rate and d-prime for the prediction based on the right amygdala ROI were (0.446±0.102), (0.356±0.073) and (0.237±0.461), correspondingly.

## Discussion

We developed a paradigm to study the behavioral and brain mechanisms that lead to long-term memory of a brief, unique experience: induced perceptual insight. We found that activity in several brain regions correlated with subsequent long-term memory of the insightful information encoded during a brief exposure to the original images (solutions) of degraded unrecognized real-world pictures (camouflages). Most notably, activity in the amygdala during the moment of induced insight was linked to long-term memory retention of the solution. Indeed, fMRI (BOLD) activity in the amygdala during exposure to the camouflage solutions could be used to predict which solutions will remain in memory a week later, on a trial-by-trial basis, with impressive reliability.

It is well established that the amygdala plays an important role in processing and encoding of emotional information (e.g., Hamann et al., 1999; McGaugh, 2004; Phelps and LeDoux, 2005), and highly emotional events are often learned in one or few trials (Rutishauser et al., 2006; Tye et al., 2008). Amygdala activity modulates the strength of emotional memories (Cahill and McGaugh, 1996), possibly by facilitating cortical processing during salient learning events (Armony et al., 1998; Paz et al., 2006). The amygdala was also shown to represent the rapidly changing value of visual stimuli that were paired with a rewarding or aversive US (Paton et al., 2006). However, the stimuli in our study did not have *a priori* emotional valence to them. None of the camouflage images in our experiments were in and of themselves emotional, and there was no pairing of those images with external rewarding or aversive events. Indeed, different subjects tended to remember different subsets of the camouflage solutions they were exposed to in a manner that was idiosyncratic and unpredictable. At the same time, the sudden realization of an insightful solution can certainly be associated with the distinct saliency of the "Aha!" moment. Our results therefore suggest that the amygdala plays a key role also in encoding events whose importance is not given *a priori* and externally, but instead is determined internally, by the organism itself. And, while better understanding of this hypothesized internal process will require further research, several observations can be made in the specific case of our study.

In the induced-insight paradigm, exposure to the solution (the original image) could lead to an abrupt and marked change in the appearance of an unchanged sensory stimulus: from a (hitherto meaningless) camouflage image to a vivid depiction of the underlying real-world scene. At the same time, we know from our pilot experiments that observers differ not only in terms of which solution images they will retain in memory days and weeks later, but also in terms of their immediate responses to the presentation of the camouflage-solution pairs. After exposure to the solution, the very same camouflage image could appear to one observer as a perceptually compelling representation of that solution, yet reported as being not very compelling by another observer -- and an opposite trend could be obtained for

another image. The ability to report the 'goodness' of the solution relies on an internal measure that may be computed automatically and involuntarily, i.e., not only when an experimenter elicits the observer's judgment. This evaluative process -- possibly mediated by the very same mechanisms that give rise to the subjective "Aha!" experience for perceptually compelling images -- may therefore be what provides the internal signal about the importance of the event, and ultimately determines how well it will be remembered. This suggests that the solutions retained in memory will be those which created a more compelling perceptual experience (a subset which differs for different subjects). The fMRI data we obtained from other regions, mainly the lateral occipital and the medial prefrontal cortices, provide converging support for this account.

The lateral occipital cortex was shown to be critical for processing of perceptual closure and surface completion, segmentation and grouping, in studies that investigated visual cortical activity before and after exposure to the solution of camouflage images (Dolan et al., 1997), as well as in studies that used other types of fragmented images of objects (e.g., Doniger et al., 2000; Grill-Spector et al., 2000; Mendola et al., 1999; Stanley and Rubin, 2003). In contrast, more anterior portions of the ventral visual pathway, and in particular the pFs, seem to be involved more in the processing of visual information about known objects. Our finding that remembered camouflage solutions are associated with increased activity in the LO, but not in the pFs, therefore suggests that the most significant changes in visual neural activity taking place during the induced insight were the reorganization of figure/ground domains and surface segmentation in the camouflage image (associated with LO), not the acquisition of information about the embedded objects (associated with the pFs). This is consistent with our proposal above, that the remembered images were those that gave rise to a more vivid perception of the underlying scene (after exposure to the solution), and also offer a concrete way for the system to evaluate the 'goodness' of a solution, by measuring the extent of neural reorganization in lateral occipital cortex.

The proposal that evaluative neural processes taking place during induced insight affect subsequent memory is supported also by the pattern of activity we observed in the medial prefrontal cortex (mPFC) and the anterior cingulate cortex (ACC). These regions have been implicated in a multitude of evaluative processes, both intentional/reflective and automatic (e.g., Amodio and Frith, 2006). In a meta-analysis of neuroimaging studies of human emotion, Phan et al. (2004) found that the mPFC was involved in nearly 50% of the studies and proposed that, taken together, the results suggest mPFC may be an integrator of affective and cognitive processing. Importantly, mPFC-amygdala interactions have also been well established in both animal and human studies (Delagado et al., 2008; Phelps et al., 2004; Quirk et al. 2003). The ACC has a well-established role in conflict monitoring and cognitive control (Botvinick et al. 2004), and it has also been proposed to take an important part in reinforcement-guided learning and representation of reward history (Rushworth et al. 2007). The ACC has also been repeatedly implicated in previous studies of insight (Subramaniam et al. 2008,; Kounios et al. 2006). In the context of our task, two main types of possibilities come to mind concerning the role of the amygdala. One is that the amygdala interacts with the integrative high-level processes in the frontal regions to evaluate the internal value expressed in the extent of the neural reorganization in visual cortex, and based on this may facilitate long-term changes in circuits, e.g. visual, that subserve the subsequent storage of the camouflage solution. Alternatively, activity in the amygdala and frontal regions may represent an evaluative process which has no causal relationship with subsequent memory. Given the known role of amygdala in memory encoding and consolidation at large (McGaugh 2004), we deem the former explanation more likely.

It is noteworthy that we did not find differential subsequent-memory-correlated activation of the hippocampal formation in our paradigm. This may result from either intensive

engagement of the hippocampal formation in non-mnemonic tasks taxed in the encoding session, or, more likely, from the possibility that whereas our memory test taps into declarative information, successful encoding in our protocol can be achieved in a non-declarative manner.

Our findings extend the known roles of amygdala in memory to include promoting of long-term memory resulting from a sudden, internal reorganization of information. The amygdala is recognized to play a crucial part in emotional learning (McGaugh, 2004; Phelps and LeDoux, 2005). Notably it also correlated with reporting insight experience in solving phrase completion task (Jung-Beeman et al. 2004), and was found to be critical for surprise-induced enhancement of learning in the rat (Holland and Gallagher 2006). Our proposal that it plays an important role in signaling to different cortical regions that an internal event of a significant neural reorganization has occurred, is consistent with these findings. What we suggest here is that amygdala influence over cortical plasticity may arise also as a result of evaluation of internal changes. The measure and benefit of the change may serve in this case as a reinforcer. This kind of mechanism may be a driving force in making cortical representations more efficient and compact.

In conclusion, we have introduced a paradigm that combines induced perceptual insight with fMRI analysis of subsequent memory performance, as a model for studying memory formation of single exposure events. We found that activity in the amygdala during the moment of induced insight could be used to predict performance in a memory task a week later, a task that required associative access to the content of the induced insight event (the pairing between a visual puzzle and its solution). We offered a framework to explain these results that also provides an integrative explanation to our other findings, of increased activity during the induced insight event in intermediate-level visual cortex (LO) and in the medial prefrontal cortex (mPFC). In this framework, the amygdala plays a central role also in memory formation of ongoing events in which the stimulus is not *a priori* emotional or externally rewarded, but leads to significant internal neural events (e.g., insight), by providing content-specific cortical regions (e.g., LO) with modulatory signals about the importance of these events.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank Merav Ahissar, Moshe Bar, Orit Furman, Efrat Furst, Kalanit Grill-Spector, Rafi Malach, Avi Mendelsohn, Morris Moscovitch, Yuval Nir, Rony Paz, Son Preminger, Robert Shapley and Nachum Ulanovsky for helpful discussions and comments on versions of the manuscript. We also thank Eunice Yang for assistance in the fMRI scans and preprocessing of fMRI data, Edna Haran-Furman for her help in the high resolution scans, Sharon Gilai-Dotan for help in delineating the LOC ROIs, and Justin Kung for help in delineating the hippocampus ROIs. This work was supported by the Minerva Foundation and the Israeli Science Foundation (YD), the National Institutes of Health grant R01EY014030 (NR) and the Weizmann Institute - NYU collaborative research fund in the neuroscience (YD and NR).

## Appendix

## Experimental Procedures

### Participants

Sixty five participants took part in this study: 37 in *Experiment 1* (ages 21-29 years, mean 24 years, 26 females), 17 in *Experiment 2* (aged 19-38, mean 25 years, 6 females), and 11 in *Experiment 3* (aged 22-29, mean 25 years, 6 females). All participants had normal or

corrected to normal vision. Participants in *Experiments 2 & 3*, which included an fMRI scan, were all right-handed. Unless otherwise indicated, participants were paid for their time.

### The Stimulus Set

The stimuli used in these experiments were 40 camouflage images that were experimentally screened out of a large collection of degraded real-world pictures that portrayed a clear, nameable object or scene. The chosen images were those in which the embedded object was not likely to be spontaneously identified, yet once the solution (the original, non-degraded image) was presented, the object embedded in the camouflage image was usually vividly perceived (i.e. the object was perceived as whole, created an impression of depth and no spurious solutions – false alarms – were perceived in the image). For the full description of the generation and pre-screening of the images see the *Generation and pre-screening of camouflage images* section in the *Supplemental Experimental Procedures*. Of the 40 images in the final set, 17 were images of animals, 8 of human figures, 3 of human faces, 3 of insects, 5 of inanimate objects and 4 contained a more complex scene that combined, for example, a human figure and an object. (For an example of an image from the set and its solution, see Figures 1 and 2.)

### Visual stimulation and response monitoring

Behavioral sessions took place in a quiet dark room, where participants were seated in front of a 19″ monitor (100Hz refresh rate). Images were presented on a medium gray background in the center of the screen and subtended a mean height of 17.5° and a mean width of 21.26° visual angle. Participants responded using the keyboard number buttons.

In the sessions performed in the MRI scanner, the visual display was fed into an LCD projector. The projected image appeared on a plastic rear-projection screen, and participants viewed the stimuli through a mirror mounted on the head coil. In *Experiment 2* the images subtended a mean height of 13.12° and a mean width of 16° visual angle; responses were collected on a 5-button RIS-418, RURB button box (Rowland Institute, Cambridge, Massachusetts). In *Experiment 3* the images subtended a mean height of 7.3° and a mean width of 10.9° visual angle; responses were collected using a response box which is part of a fORP system that includes 8 button handhelds response box manufactured by Current Designs Inc. (Philadelphia PA, USA).

Experiments were programmed in Presentation 0.81 (Neurobehavioral Systems Inc. www.neurobs.com). Participants were briefed about the task with written instructions and examples that were presented as a slide show, and performed several practice trials until they understood the task. The stimuli used in the practice trials were not drawn from the set of 40 images used in the experiment itself.

### Experimental Protocol

Three experiments are reported, *Experiment 1, 2* and *3*. The overall protocol in all the experiments was similar and consisted of two sessions each, *'study'* and *'test'*. *Experiment 1* was behavioral only, and was conducted to determine memory performance over time and select the time interval between the *'study'* and *'test'* sessions to use in Experiments 2 and 3. In *Experiment 1*, separate groups of participants performed the *'test'* session 15 minutes, 24 hours, 1 week and 3 weeks after the *'study'* session (9-10 participants in each group). In Experiments 2 and 3 the *'study'* session was performed while participants were undergoing brain imaging in the fMRI scanner. The *'study'* protocol was therefore slightly modified from *Experiment 1* to adapt it to the fMRI environment. The protocol described below (Figure 3A) is that of *Experiment 2*. (For descriptions of the slightly different *'study'* session protocols in Experiments 1 and 3 see Supplemental Figs. S1 and S3 and *Supplemental*

*Experimental Procedures*). The *'test'* session in *Experiment 2* and *Experiment 3* was identical to that of *Experiment 1* (Figure 3B) and performed 1 week after *'study'*.

In the *'study'* session, 30 camouflage images were presented, chosen randomly for each participant out of the set of 40 images (in *Experiment 3*, 40 images were presented). Each camouflage image was presented for 10 sec (*CAM1*). Participants were instructed to press a button if they thought they recognized the underlying scene during the presentation of the camouflage in *CAM1* (the image remained on the screen for 10 sec regardless of whether and when the participant pressed the button). Note that the indication of recognition at this stage is not necessarily accurate (it may include false-alarms or exclude correct recognitions in which the participant is not sure). *CAM1* was followed by 4 sec in which the solution (the original gray-level image) and the camouflage alternated four times, each presented for half a sec (*SOL*). Next, participants were presented again with the camouflage for two sec (*CAM2*). Finally, to assess spontaneous recognition, a question appeared: "Did you identify the object in the camouflage image before the solution?" (*QUERY*). Participants were instructed to answer 'yes' even if they only partially recognized the scene, as long as they discerned the main object. They were also instructed to answer 'yes' if they recognized the object during *CAM1* even if they did not press the button at that stage (e.g. because they were not sure their perception was correct) and, conversely, to answer 'no' even if they indicated recognition during *CAM1* in cases when during *SOL* they saw that the camouflaged object was different from what they thought they recognized during *CAM1*.

To allow the separation of the signal evoked by each event, a blank screen (medium gray) was presented as an inter-stimulus interval (ISI) for a period of 2, 3 or 4 sec (picked randomly) between the *CAM1, SOL* and *CAM2* stages. A randomly picked inter-trial interval (ITI) of 3, 4 or 5 sec was also used.

For determination of memory performance in the *'test'* session (Figure 3B), all 40 camouflage images from the stimulus set were presented to each participant. Thirty images (and their solutions) were therefore previously seen (in the *'study'* session), and the remaining 10 images were novel (different ones for different participants). Each camouflage image was presented for 10 sec, and the participants were instructed to press a button if they thought they identified the object in the image. Once they pressed the button, or after 10 sec if they did not press it, the four-option Multiple-Choice question appeared, and the participants were requested to identify the object in the image by pressing one of four numbers on a keypad. The question remained on the screen until a response was given, but not for longer than 10 sec, and the next camouflage image was then presented. If a choice was made within the 10 sec limit, participants were next asked to indicate their level of confidence: 'guess', 'fairly confident', or 'completely confident'. Regardless of the confidence level indicated, if their answer to the Multiple-Choice question was correct, the camouflage image was re-presented, with a Grid-map of numerals (1-9) superimposed on it (Grid task). Participants were asked to specify the Grid-map numeral overlaid on a specific feature of the object (e.g., 'the eye of the clown', Figure 3B). The object features queried at the Grid questions were ones that are not likely to be distinctive by themselves, but rather those which their identification was facilitated by a holistic perception of the embedded object. The image with the Grid-map was presented until a response was provided but not for more than 10 sec. The chance level, combining the Multiple-Choice question and the Grid task question, is only 2.75% (25% at the multiple choice question followed by 11% at the grid-map question). Therefore, in all subsequent analyses correct responses on both tests were used as indication that the observer perceived the underlying scene correctly.

This classification was supported by two analyses: First, we compared the number of these images with the expected number of chance correct answers in the Multiple-Choice task.

The expected number of chance correct answers (computed individually for each participant) is the number of images that were indicated as not identified in the *'study'* and that their perceptual identification was not verified in the Grid task, divided by four. The comparison of those two sets of numbers revealed that they did not differ significantly ($p=0.15$). Second, inspection of the participants' confidence report revealed that they are much less confident in their answers for these images, than for images that were answered correctly in the Grid task. For correct Multiple-Choice answers, participants indicate that they are highly confident in 62% of the Grid correct images (79 out of 128) but only in 37% of the Grid wrong images (14 out of 38). This is even more marked for the Grid task confidence, where they are highly confident in 77% of the Grid correct images but only in 19% of the Grid wrong images. Our conclusion was that images for which participants did not provide a correct answer to the Grid task should be considered as not having been retained in memory, or retained only semantically.

Statistical analysis of the behavioral data described in the *Results* section was done using Statistica (StatSoft, Inc., 2004; version 6; www.statsoft.com).

## Functional neuroimaging

MRI scanning during the *'study'* session of *Experiment 2* was conducted on a 3 Tesla head-only Siemens Allegra scanner at the Center for Brain Imaging (CBI) in New York University. Seventeen healthy participants took part in the imaging experiment. Thirteen of them were paid for their participation. Informed consent was obtained from all participants, and all procedures were approved by the New York University Committee on Activities Involving Human Subjects. Three participants were omitted from the analysis, one because of excessive movements in the magnet and two because they did not complete the *'test'* session.

Structural scans (T1-weighted) were obtained with a head coil (transmitter/receiver; Nova Medical, Wakefield MA, model NM011). Functional scans used the same head coil for excitation (transmit) and a flexible four element array of surface coils placed evenly around the head for detection (receive; Nova Medical, Wakefield MA, model NMSC011). Two types of high-resolution T1-weighted scans were obtained for each participant: (i) a set taken with an MPRAGE sequence resulting in 1×1×1 mm voxels (256*256); (ii) a set acquired with a T1-weighted spin echo sequence resulting in 3×1.5×1.5 mm voxels (128*128), taken with the same slice prescription as that used for the functional scans (see below); the scan was used to obtain a precise alignment between the functional data and the high resolution MPRAGE images. Functional (T2*-weighted) EPI images (TR=2 sec, TE=30ms, flip angle=90°) were acquired with an in-plane resolution of 64×64 resulting in 3×3×3 mm voxels.

**Scanning during the 'study' session—**In *Experiment 2*, participants were continually scanned during presentation of the 30 camouflage images of the *'study'* session. Each trial lasted 20-34 sec, separated by an inter-trial interval (ITI) of 3-5 sec. The *'study'* scans lasted a total of 775-809 sec.

**Lateral-occipital region-of-interest 'localizer' runs—**After completion of the *'study'* session in *Experiment 2*, each participant performed another functional run, whose aim was to localize the lateral occipital complex (LOC; Grill-Spector et al., 1998; Kanwisher et al., 1996; Malach et al., 1995). This run consisted of 16 sec blocks of pictures of everyday objects alternating with blocks of scrambled versions of the same objects. The two types of blocks repeated 12 times for a total of 384 sec (192 fMRI acquisitions). The images used are described in Stanley & Rubin 2003. They were shown in isolation on a homogeneous

background, and had a mean height of 7.6°, and mean width of 8.9°. Successive images were jittered +/-0.6°. Participants were required to maintain fixation, observe the images and press a button whenever the same image repeated twice consecutively.

MRI scanning during the *'study'* session of *Experiment 3* was conducted on a 3T Trio Magnetom Siemens scanner at the Weizmann Institute of Science. Eleven healthy participants took part in the imaging experiment. They were all paid for their participation. Informed consent was obtained from all participants. Two participants were discarded from the analysis since they had almost no *REM* trials (one participant had one *REM* trial and the other had two trials), and hence their data could not be used for subsequent memory prediction.

All images were acquired using a 12-channel head matrix coil. Three-dimensional T1-weighted anatomical scans were acquired with high resolution 1-mm slice thickness (3D MP-RAGE sequence, TR 2300ms, TE 2.98ms, $1mm^3$ voxels). Functional high resolution scans were acquired resulting in a 2*2*2 mm voxels (22 slices without gap, TR=2000ms, TE=36ms, Flip angle 75°). The slices were obtained at 30° towards the coronal plane from ACPC, with the amygdala at the center of the FOV, covering also most of the Hippocampus, most of the Temporal lobes & the inferior half of the Frontal lobes (see Supplemental Figure S3). To obtain a precise alignment between the functional data and the MPRAGE images, a T1-weighted spin echo sequence resulting in 2×1×1 mm voxels was taken with the same slice prescription as that used for the functional scans.

**Scanning during the 'study' session—**In *Experiment 3*, participants were continually scanned during presentation of the 40 camouflage images of the *'study'* session. Each trial lasted 22-38 sec, separated by an inter-trial interval (ITI) of 4-8 sec. The *'study'* scans lasted a total of 1358-1416 sec.

## fMRI data analysis

Unless otherwise indicated, fMRI data were processed using the BrainVoyager QX 1.3 software package (Brain Innovation, Maastricht, Netherlands). Data were first corrected for head motion (scans with head movement larger than 2 mm were rejected) and for slice-timing acquisition. The runs were high-pass filtered according to the period of stimulation (at 0.016 Hz for the camouflage runs and at 0.005 for the localizer runs). The complete data set was converted into Talairach space. For the multi-subject voxel-by-voxel GLM analyses (see below), data from the camouflage runs were spatially smoothed with a 6 mm (full-width at half-height) Gaussian kernel. In all other analyses, which were subject-specific, data were not spatially smoothed.

**General linear model (GLM) Experiment 2—**Data from the camouflage *'study'* runs were modeled by assigning a predictor to each trial based on the behavioral performance: *SPONT*, for those trials when the camouflage was reported (at the *QUERY* stage) as identified spontaneously during the *'study'* session (i.e. trials in which the solution served only as a confirmation of correct spontaneous perception and not as an event of perceptual insight), *REM* ('remembered'), trials in which the camouflage image was not spontaneously identified during *CAM1* and the solution was subsequently remembered, yielding correct performance on both the Multiple-Choice and the Grid tasks at the *'test'* a week later (i.e. trials in which the solution served as a learning event); *NotREM*, ('not remembered'), trials in which the camouflage was not identified during the *'study'* session and its solution was not remembered during the *'test'*. Each of the three protocol stages (*CAM1*, *SOL*, *CAM2*) was assigned a separate predictor. Combined with the labels of performance, this resulted in nine predictors (*CAM1-REM*, *CAM1-NotREM*, *CAM1-SPONT* etc.). An additional predictor,

blank, was used for all the time frames in which the participants viewed a gray screen. These include 10 sec prior to the start of the camouflage run, and the ISI's and ITI's during the run. For each predictor, a box-car function valued 1 (and 0 for the blank predictor) was convolved with a canonical hemodynamic response function (Boynton et al. 1996).

For each comparison of interest contrasts were created between the appropriate predictors (the main contrast compared activation during *REM* and *NotREM* trials, and see also 'Regions of interest', and in Results), and p-values were calculated (t-test) for each voxel. For the *SOL* vs. *Baseline* and the object localizer 'objects' vs. 'scarmbled-objects' contrasts, the p-values were adjusted for multiple comparisons using False Discovery Rate controlling procedures (Benjamini and Hochberg, 1995; Genovese et al., 2002; Stanley and Rubin, 2003) before thresholding. Finally, voxels that did not belong to contiguous clusters of at least 5 significant functional voxels were eliminated.

**Regions of interest (ROIs) Experiment 2—**These were defined in three different ways. First, and based on prior results indicating the occipito-temporal stream as crucial to shape perception and object recognition (Grill-Spector & Malach 2004), visual cortical ROIs were created from the data obtained in the localizer scan. Data from those runs were modeled using a boxcar predictor for each experimental condition except fixation ('objects' and 'scrambled objects'). A hemodynamic lag of 4 or 6 sec was fitted to the model of each subject by maximizing the extent of the overall visual activations. Statistical maps were created, separately for each observer, by contrasting the 'objects' and 'scrambled objects' predictors and thresholded at $q<0.005$. ROIs were delineated from the resulting map in three separate portions of occipito-temporal cortex (as identified by anatomical markers; Duvernoy 1999): the lateral occipital sulcus (LO), the posterior to mid parts of the fusiform gyrus (pFs; Grill-Spector et al., 2000) and the CoS. In addition, voxels which were significantly more active during the 'scrambled objects' condition were delineated in and around the calcarine fissure as an early visual ROI (EarlyVis; Supplemental Material, Figure S2). We were able to delineate the LO and EarlyVis ROIs in all 14 participants, the pFus ROI in 12 participants, the left CoS in 10 participants, and the right CoS in 8 participants.

Second, another set of ROIs was generated from the data obtained in the camouflage run itself by contrasting activity during the *SOL* stage of all three event types (*SPONT*, *REM* and *NotREM*) with activity during the time period of 'baseline' (blank) trials and thresholding at $q<0.05$. Note that by collapsing across all event types, the resulting statistical map, and the ROIs extracted from it, were unbiased with respect to subsequent memory performance (Kensinger and Corkin, 2004). This contrast resulted in extensive activations in visual as well as frontal areas, and also in prominent activation clusters in bilateral amygdala (for the full list of activations see Supplemental Material, Table S1). To examine more closely the activity in those brain regions that were particularly engaged during the presentation of the solution (and for which we did not have independent localizer data), we delineated from the results of this contrast ROIs in the frontal cortex (in the lateral orbital gyrus and in the inferior frontal sulcus) and in the amygdala (as identified by anatomical markers; Duvernoy, 1999). Following Johnstone et al. (2005), the amygdala ROI was defined separately for each subject. For each participant we took the clusters of 3 contiguous functional voxels activated in conjunction with group level activation. We were able to delineate activation in the left amygdala for 11 participants, but only for 6 in the right amygdala (even when the threshold for the right amygdala was relaxed to $q<0.15$ we were able to delineate the right amygdala ROI only for 8 subject, and the additional data did not change the results).

Finally, we delineated the hippocampus based on the high-resolution T1-weighted MRI images and on established anatomical landmarks (Pruessner et al., 2000). Three separate hippocampal ROIs were defined, for each observer separately: head, body and tail.

For each of the ROIs we extracted the time-course obtained during the camouflage *'study'* session, separately in each participant. The time-courses were linearly interpolated from the TR resolution (2 sec) to 1 sec resolution, to fit the protocol time-course, and transformed into percent-signal-change, based on the two TRs preceding each event. (This and all other time-course calculations were done using Matlab, The Math-works, Inc., Natick, MA, version 6.1, 2001.)

*ROI time-course data* were first analyzed by computing event-triggered averages for each event type (*SPONT, REM* and *NotREM*), separately for each stage in the trial (*CAM1, SOL, CAM2*; Figure 3A). These curves were calculated for each ROI (with means and standard errors computed across participants), and allowed us to visually inspect the shapes of the hemodynamic response functions.

To perform statistical testing on the ROI time-course data, the BOLD activity values from 4 to 10 sec after the onset of each stage of the protocol were extracted for each trial (or 4 to 12 sec for *CAM1* events, which were longer). These corresponded to the peak time points of the hemodynamic response functions obtained for each stage from the event-triggered averages. Each series of time points was labeled with the behavioral performance associated with it (*SPONT, REM* or *NotREM*) and with the participant's index. In this manner we obtained for each stage of the trial (*CAM1, SOL, CAM2*) a matrix of 420 rows (30 trials per participant * 14 participants) and 9 columns (7 time points, plus one column for participant index and one for behavioral performance status, i.e. event type, 11 columns for *CAM1* events). Such a matrix was obtained for each ROI (for those ROIs that were identified in less than the full set of 14 participants the number of rows was accordingly smaller). The matrices were imported into Statistica (Statsoft inc.) and the values at each time point were subjected to a mixed-model ANOVA with event type (*REM, NotREM, SPONT*) as one factor, and participant index as a random factor, to determine whether there were significant differences between the BOLD activity within the same region in different event types. A comparison between two event types was considered significant if the resulting p-value for 3 consecutive time points was significant, with a criterion $\alpha = 0.05$ and a Bonferroni correction for multiple comparisons (for *SOL* and *CAM2* 3 consecutive time points out of 7 time points provide 5 comparisons, therefore $\alpha^* = 0.01$; for *CAM1* 3 consecutive time points out of 9 time points provide 7 comparisons, therefore $\alpha^* = 0.007$).

Among the trials where participants did not spontaneously recognize the image during *'study'*, the total number of trials that were performed correctly during *'test'* ('remembered', or *REM* events) was 128 (across all participants). The total number of trials where an error was performed during *'test'* (in the Multiple Choice, Grid task, or both; *NotREM* events) was 178. The number of trials where the underlying object was recognized spontaneously during *CAM1* (*SPONT* events) was 114.

**General linear model (GLM) Experiment 3—**Data from the camouflage *'study'* runs of *Experiment 3* were modeled in the same manner as in *Experiment 2*, except that to avoid circularity of the prediction the subsequent memory information was not used in the GLM. The predictors were hence: *SPONT*, for those trials when the camouflage was reported (at the *QUERY* stage) as identified spontaneously during the *'study'* session (i.e. trials in which the solution served only as a confirmation of correct spontaneous perception and not as an event of perceptual insight), and *NotIdentified*, for those trials when the camouflage was reported as not identified spontaneously during the *'study'* session.

**Regions of interest (ROIs) Experiment 3—**The only ROIs delineated from the data of *Experiment 3* were those of the amygdala, for the purpose of subsequent memory trial-by-trial prediction. The amygdala ROIs were generated similarly to *Experiment 2*, this time by

contrasting activity during the *SOL* stage of both event types (*SPONT* and *NotIdentified*) with activity during the time period of 'baseline' (blank) trials. In this dataset we were able to delineate activation in the left amygdala for 8 of the 9 participants, and again only for 6 in the right amygdala.

**Trial-by-trial subsequent memory prediction Experiment 3—**For each of the amygdala ROIs we extracted the time-course obtained during the camouflage *'study'* session, separately in each participant. The time-courses were linearly interpolated from the TR resolution (2 sec) to 1 sec resolution, to fit the protocol time-course, and transformed into percent-signal-change, based on the two TRs preceding each event. The BOLD activity values from 6 to 14 sec after the onset of the *SOL* stage of the protocol were extracted for each trial, and the area under the curve was computed. Each series of time points was labeled with the behavioral performance associated with it (*SPONT* or *NotIdentified*) and with the participant's index. Next, we sorted the *NotIdentified* trials by the area under the curve value. At *Experiment 1* and *Experiment 2 REM* trials consisted on average 40% of the total number of the *NotIdentified* images. Hence we divided the sorted trials list into the top 40% which were labeled *Predicted-REM*, and the bottom 60% which were labeled *Predicted-notREM*. For each subject we then computed the hit and false alarm rate of the prediction, as compared with the actual subsequent memory performance. (This was done using Matlab, The Math-works, Inc., Natick, MA, version 6.1, 2001.)

## References


Amodio DM, Frith CD. Meeting of minds: the medial frontal cortex and social cognition. Nature Rev. Neurosci. 2006; 7:268–277. [PubMed: 16552413]

Armony JL, Quirk GJ, LeDoux JE. Differential effects of amygdala lesions on early and late plastic components of auditory cortex spike trains during fear conditioning. J Neurosci. 1998; 18:2592–2601. [PubMed: 9502818]

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Statist Soc Ser B. 1995; 57:289–300.

Botvinick MM, Cohen JD, Carter CS. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn. Sci. 2004; 8:539–546. [PubMed: 15556023]

Bowden EM, Jung-Beeman M, Fleck J, Kounios J. New approaches to demystifying insight. Trends Cog. Sci. 2005; 9:322–328.

Boynton GM, Engel SA, Glover GH, Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci. 1996; 16:4207–4221. [PubMed: 8753882]

Cahill L, McGaugh JL. Modulation of memory storage. Curr Opin Neurobiol. 1996; 6:237–242. [PubMed: 8725966]

Brewer JB, Zhao Z, Desmond JE, Glover GH, Gabrieli JD. Making memories: brain activity that predicts how well visual experience will be remembered. Science. 1998; 281:1185–1187. nbsp, and E. [PubMed: 9712581]

Cahill L, McGaugh JL. Modulation of memory storage. Curr Opin Neurobiol. 1996; 6:237–242. [PubMed: 8725966]

Delgado MR, Nearing KI, LeDoux JE, Phelps EA. Neural circuitry underlying the regulation of conditioned fear and its relation to extinction. Neuron. 2008; 59:829–838. [PubMed: 18786365]

Dolan RJ, Fink GR, Rolls E, Booth M, Holmes A, Frackowiak RSJ, Friston KJ. How the brain learns to see objects and faces in an impoverished context. Nature. 1997; 389:596–599. [PubMed: 9335498]

Doniger GM, Foxe JJ, Murray MM, Higgins BA, Snodgrass JG, Schroeder CE, Javitt DC. Activation time course of ventral visual stream object-recognition areas: high density electrical mapping of perceptual closure processes. J. Cogn. Neurosci. 2000; 12:615–621. [PubMed: 10936914]

Dudai Y. The neurobiology of consolidations, or, how stable is the engram. Annu Rev Psychol. 2004; 55:51–86. [PubMed: 14744210]

Duvernoy, HM. The Human Brain: Surface Three-Dimensional Section Anatomy and MRI. 2nd ed. Springer; New York: 1999.

Gauthier I, Tarr MJ. Becoming a “Greeble” expert: exploring mechanisms for face recognition. Vision Res. 1997; 37:1673–1682. [PubMed: 9231232]

Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage. 2002; 15:870–878. [PubMed: 11906227]

Grill-Spector K, Malach R. The human visual cortex. Ann Rev Neurosci. 2004; 27:649–677. [PubMed: 15217346]

Grill-Spector K, Kushnir T, Hendler T, Malach R. The dynamics of object-selective activation correlate with recognition performance in humans. Nat Neurosci. 2000; 3:837–843. [PubMed: 10903579]

Grill-Spector K, Kushnir T, Hendler T, Edelman S, Itzchak Y, Malach R. A sequence of object-processing stages revealed by fMRI in the human occipital lobe. Hum Brain Map. 1998; 6:316–328.

Hamann SB, Ely TD, Grafton ST, Kilts CD. Amygdala activity related to enhanced memory for pleasant and aversive stimuli. 1999; 2:289–293.

Hasson U, Furman O, Clark D, Dudai Y, Davachi L. Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. Neuron. 2008; 57:452–462. [PubMed: 18255037]

Hebb, D. The Organization of Behavior: A Neuropsychological Theory. Wiley; NY: 1949.

Holland PC, Gallagher M. Different roles for amygdala central nucleus and substantia innominata in the surprise-induced enhancement of learning. J. Neurosci. 2006; 26:3791–3797. [PubMed: 16597732]

Johnstone T, Somerville LH, Alexander AL, Oakes TR, Davidson RJ, Kalin NH, Whalen PJ. Stability of amygdala BOLD response to fearful faces over multiple scan sessions. NeuroImage. 2005; 25:1112–1123. [PubMed: 15850729]

Jung-Beeman M, Bowden EM, Haberman J, Frymiare JL, Arambel-Liu S, Greenblatt R, Reber PJ, Kounios J. Neural activity when people solve verbal problems with insight. PLoS Biol. 2004; 2:E97. [PubMed: 15094802]

Kanwisher N, Chun MM, McDermott J, Ledden PJ. Functional imagining of human visual recognition. Cogn Brain Res. 1996; 5:55–67.

Kaplan C, Simon H. In search of insight. Cog Psychol. 1990; 22:374–419.

Karni A, Sagi D. Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. Proc Natl Acad Sci USA. 1991; 88:4966–4970. [PubMed: 2052578]

Kensinger EA, Corkin S. Two routes to emotional memory: Distinct neural processes for valence and arousal Proc Natl Acad Sci USA. 2004; 101:3310–3315.

Köhler, W. The Mentality of Apes. Routledge and Kegan Paul Ltd; London: 1925.

Kounios J, Frymiare JL, Bowden EM, Fleck JI, Subramaniam K, Parrish TB, Jung-Beeman M. The prepared mind: neural activity prior to problem presentation predicts subsequent solution by sudden Insight. Psychol Sci. 2006; 17:882–890. [PubMed: 17100789]

Kounios J, Fleck JI, Green DL, Payne L, Stevenson JL, Bowden EM. The origins of insight in resting-state brain activity. Neuropsychologia. 2008; 46:281–291. [PubMed: 17765273]

Maier N. Reasoning in humans II: the solution of a problem and its appearance in consciousness. J Comp Psychol. 1931; 12:181–194.

Malach R, Reppas J, Benson R, Kwong K, Jiang H, Kennedy W, Ledden P, Brady T, Rosen B, Tootell R. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci USA. 1995; 92:8135–8139. [PubMed: 7667258]

Martin SJ, Grimwood PD, Morris RGM. Synaptic plasticity and memory: an evaluation of the hypothesis. Ann. Rev. Neurosci. 2000; 23:649–711. [PubMed: 10845078]

McGaugh JL. The amygdala modulates the consolidation of memories of emotionally arousing experiences. Ann. Rev. Neurosci. 2004; 27:1–28. [PubMed: 15217324]

Mendola JD, Dale AM, Fischl B, Liu AK, Tootell RBH. The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. J. Neurosci. 1999; 19:8560–8572. [PubMed: 10493756]

Newell, A.; Rosenbloom, PS. Mechanisms of skill acquisition and the law of practice. In: Anderson, JR., editor. Cognitive skills and their acquisition. Lawrence Erlbaum Associates; New Jersey: 1981. p. 1-55.

Paller K, Kutas M, Mayes A. Neural correlates of encoding in an incidental learning paradigm. Electroencephalogr Clin Neurophysiol. 1987; 67:360–371. [PubMed: 2441971]

Paton JJ, Belova MA, Morrison SE, Salzman CD. The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature. 2006; 439:865–870. [PubMed: 16482160]

Paz R, Pelletier JG, Bauer EP, Paré D. Emotional enhancement of memory via amygdala-driven facilitation of rhinal interactions. Nat Neurosci. 2006; 9:1321–1329. [PubMed: 16964249]

Phan KL, Wager TD, Taylor SF, Liberzon I. Functional neuroimaging studies of human emotions. CNS Spectrums. 2004; 9(4):258–266. [PubMed: 15048050]

Phelps EA, LeDoux JE. Contributions of the amygdala to emotion processing: from animal models to human behavior. Neuron. 2005; 48:175–187. [PubMed: 16242399]

Phelps EA, Delgado MR, Nearning KI, LeDoux JE. Extinction learning in humans: role of the amygdala and vmPFC. Neuron. 2004; 43:897–905. [PubMed: 15363399]

Porter P. Another picture-puzzle. Am J Psych. 1954; 67:550–551.

Pruessner JC, Li LM, Serles W, Pruessner M, Collins DL, Kabani N, Lupien S, Evans AC. Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: Minimizing the Discrepancies between Laboratories. Cereb. Cortex. 2000; 10:433–442. [PubMed: 10769253]

Quirk GJ, Likhtik E, Pelletier JG, Paré D. Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. J Neurosci. 2003; 23:8800–8807. [PubMed: 14507980]

Roediger, HL.; Dudai, Y.; Fitzpatrick, SM., editors. Science of memory: concepts. Oxford University Press; New-York: 2007.

Rubin N, Nakayama K, Shapley R. Abrupt learning and retinal size specificity in illusory-contour perception. Current Biology. 1997; 7:461–467. [PubMed: 9210373]

Rubin, N.; Nakayama, K.; Shapley, R. The role of insight in perceptual learning: Evidence from illusory contour perception. In: Fahle, M.; Poggio, T., editors. Perceptual Learning. MIT Press; Cambridge MA: 2002. p. 235-252.

Rushworth MF, Buckley MJ, Behrens TE, Walton ME, Bannerman DM. Functional organization of the medial frontal cortex. Curr Opin Neurobiol. 2007; 17:220–227. [PubMed: 17350820]

Rutishauser U, Mamelak AN, Schuman EN. Single-trial learning of novel stimuli by individual neurons of the human hippocampus-amygdala complex. Neuron. 2006; 49:805–813. [PubMed: 16543129]

Seitz AR, Watanabe T. The phenomenon of task-irrelevant perceptual learning. Vis Res. 2009; 49:2604–2610. [PubMed: 19665471]

Squire, LR.; Kandel, ER. Memory: From Mind to Molecules. Scientific American Books; NY: 1999.

Stanley DA, Rubin N. fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. Neuron. 2003; 37:323–331. [PubMed: 12546826]

Sternberg, R.; Davidson, J. The nature of insight. Cambridge University Press; Cambridge, MA: 1995.

Stevens JC, Savin HB. On the form of learning curves. J Exp Anal Behav. 1962; 5:15–18. [PubMed: 13917061]

Subramaniam K, Kounios J, Parrish TB, Jung-Beeman M. A brain mechanism for facilitation of insight by positive affect. J Cogn Neurosci. 2009; 21:415–432. [PubMed: 18578603]

Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, Witter MP, Morris RGM. Schemas and memory consolidation. Science. 2007; 316:76–82. [PubMed: 17412951]

Tulving, E. Elements of episodic memory. Oxford University Press; Oxford: 1983.

Tye KM, Stuber GD, de Ridder B, Bonci A, Janak PH. Rapid strengthening of thalamo-amygdala synapses mediates cue-reward learning. Nature. 2008; 453:1253–1257. [PubMed: 18469802]

Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL. Building Memories: Remembering and Forgetting of Verbal Experiences as Predicted by Brain Activity. Science. 1998; 281:1188–1191. [PubMed: 9712582]

**Figure 1.**
Example of a degraded real-world picture, or ‘camouflage’ image, used in the study. (For the original, ‘solution’ image see Figure 2 overleaf.)

**Figure 2.**
The original picture, or ‘solution’ image used to generate Figure 1.

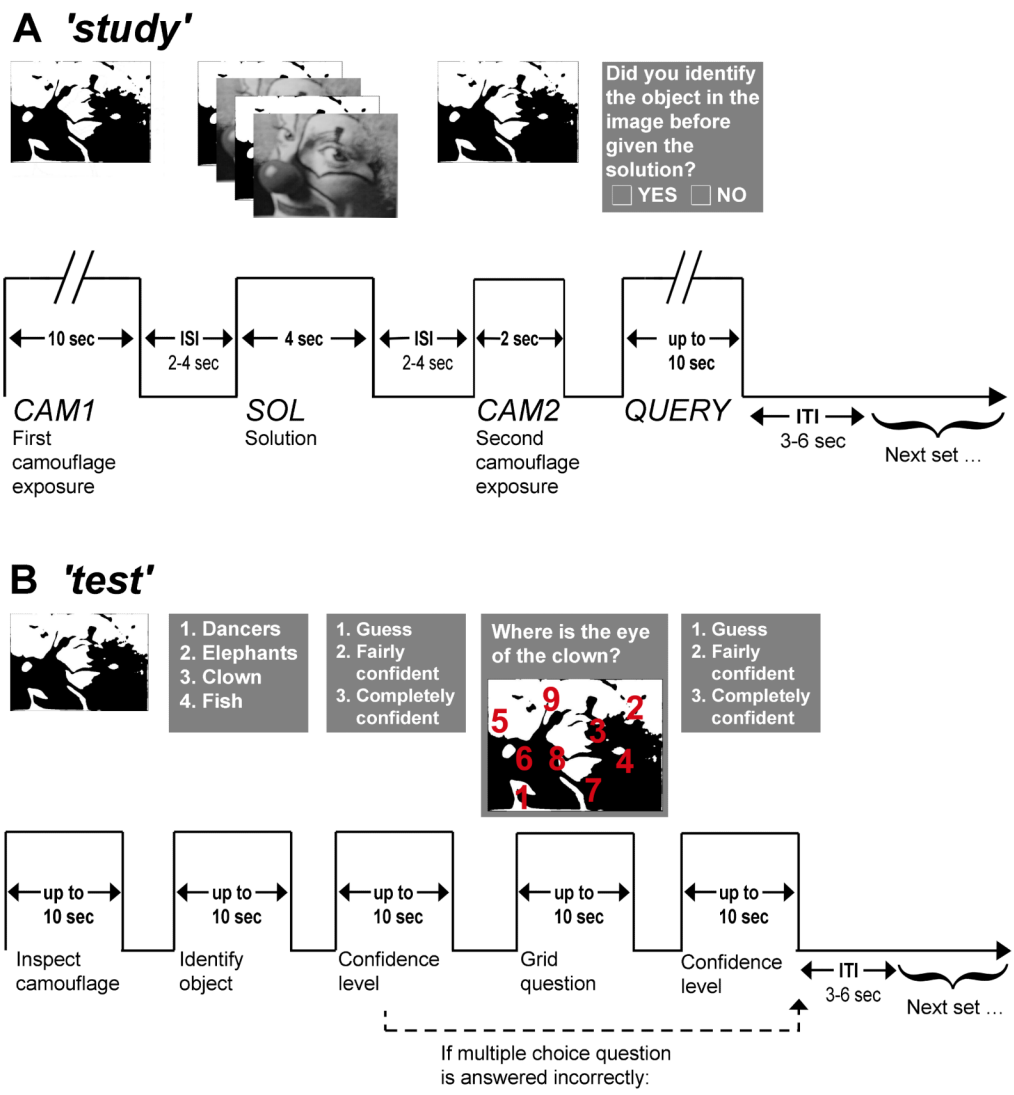


**Figure 3.**

The protocol of *Experiment 2*. **A.** *'study'* session, performed in the fMRI scanner. The session included 30 trials. Each consisted of: *CAM1*, the first camouflage exposure (participants press 'YES' if they think they identified the hidden object); *ISI* (inter stimulus interval, blank screen); *SOL*, solution (camouflage and original images alternate at 0.5 sec epochs); *ISI*; *CAM2*, a second camouflage exposure; *QUERY*, participants report their recognition at *CAM1*. Trials lasted 20-30 secs each and were separated by an intertrial intervals (*ITI*) of 3-6 secs. **B.** *'test'* session, performed 7 days after the *'study'*, outside the scanner. This session included 40 trials: the 30 camouflage images presented in the *'study'* session, interleaved with 10 novel images. In each trial, participants were given up to 10 secs to press a button if they recognized the camouflage image. They were then given a Multiple Choice recognition test (four alternatives), followed by a confidence rating (three levels). If they answered the Multiple Choice correctly, participants were presented with the camouflage image superimposed with a grid of numerals (Grid task), and instructions to find the numeral at the location of a specific feature in the image.

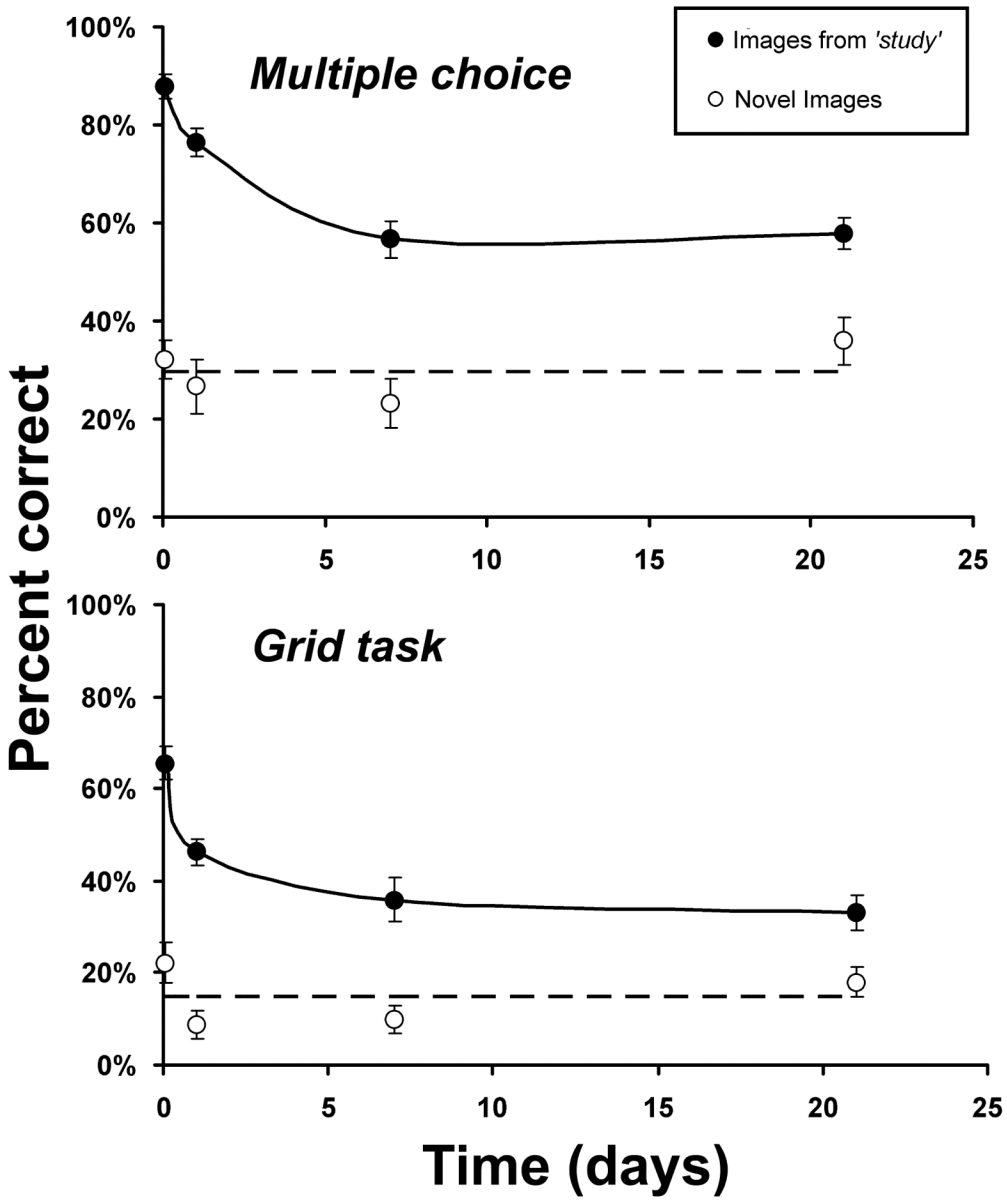


**Figure 4.**

Forgetting curves in Experiment 1. *Top panel*, performance on the Multiple Choice task by separate groups of participants tested 15 minutes, 1 day, 7 days and 21 days after a *'study'* session, respectively. *Bottom panel*, performance of the same groups on the subsequent Grid task. In both panels, images which participants recognized spontaneously during *'study'* were excluded from the calculation of the percentages correct. The dashed lines depict average performance of all the groups on the 10 images not presented during *'study'*; these novel images were a different, randomly selected subset for each observer. For the slightly different *'study'* protocol of *Experiment 1* see Supplemental Figure S1.

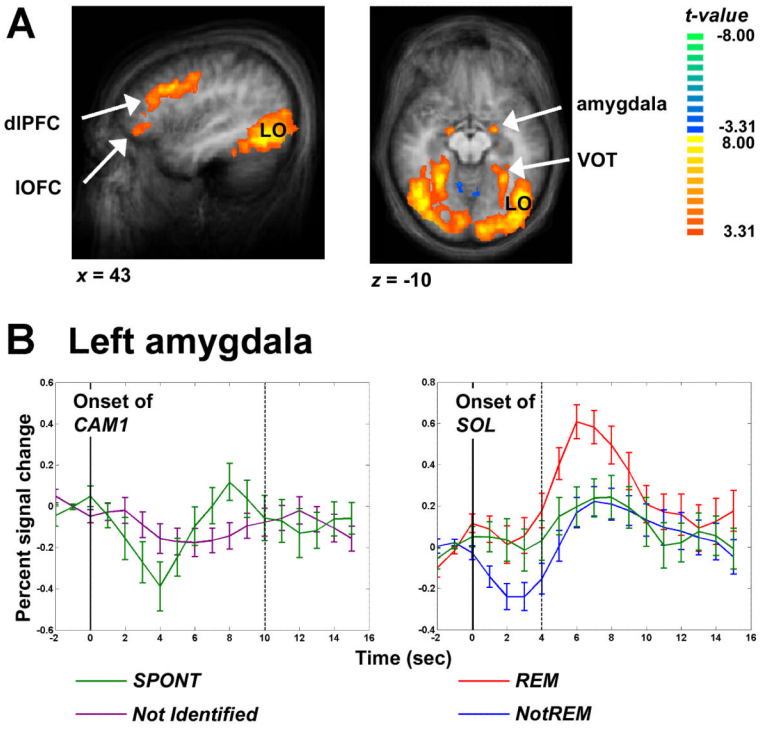


**Figure 5.**

During presentation of the solution images (SOL), the left amygdala was the brain region showing the greatest difference in activity between subsequently remembered and not remembered images (*REM* vs. *NotREM*). **A.** Sagittal and axial views of brain regions obtained by contrasting activation during SOL, collapsed across all trial types, with the baseline condition (blank screen). This analysis, which is unbiased with respect to subsequent memory or spontaneous recognition performance, was used to delineate the amygdala ROI. Other regions showing *SOL*-related activity are also shown (LO, lateral occipital; VOT, ventral occipito-temporal; dlPFC, dorso-lateral prefrontal cortex; lOFC, lateral orbital frontal cortex. For the full list see Supplemental Table S1.) **B.** Time-courses of activity during presentation of the camouflage image and its solution in the left amygdala. *Left panel:* percent-signal-change average during *CAM1*, sorted according to recognition performance as reported in the *'study' QUERY* stage. Green, activity while observing spontaneously identified images (*SPONT*); purple, Not identified images (*REM* and *NotREM* collapsed). *Right panel:* percent-signal-change average during *SOL*. sorted according to subsequent memory performance in *'study'*. red, subsequently remembered (*REM*); blue, not remembered (*NotREM*). (The green curve again denotes activity while observing spontaneously identified images, *SPONT*).

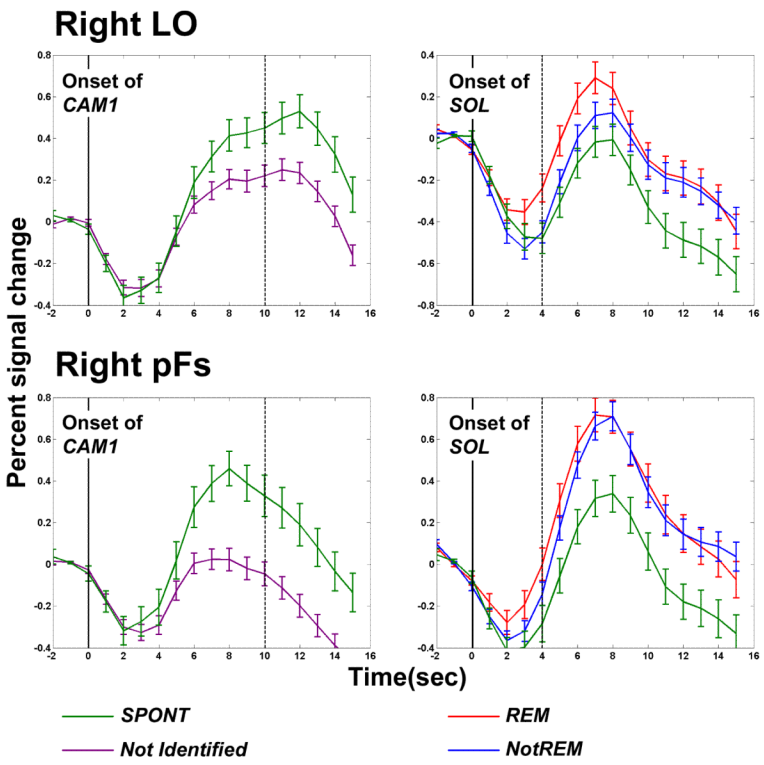


**Figure 6.**

Spontaneous identification and subsequent memory effects in mid-level visual cortex. *Left panels*: event-triggered average activity during *CAM1*, sorted according to recognition performance as reported in the *'study' QUERY* stage. Green, activity while observing spontaneously identified images (*SPONT*); Purple, *Not identified* images. *Right panels:* averaged activity during *SOL*, sorted according to subsequent memory performance in *'study'*. red, subsequently remembered (*REM*); blue, not remembered (NotREM); green, spontaneously identified images (*SPONT*). The ROIs were delineated based on the visual localizer runs (see Experimental Procedures). LOC, lateral occipital complex; pFs, posterior fusiform sulcus. Time-courses presented here were extracted from the right hemisphere ROIs. For a view of these ROIs see Supplemental Figure S2.

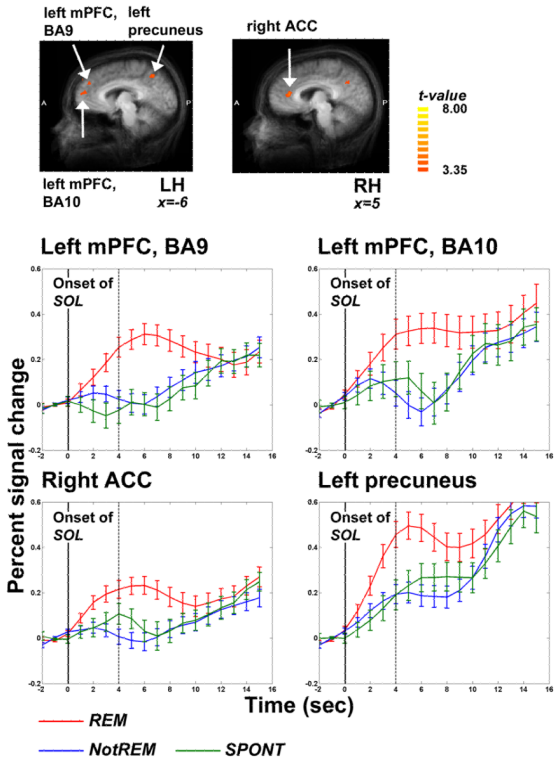


**Figure 7.**

Brain regions showing higher activity during presentation of the solutions of subsequently remembered camouflage images. *Top Panels*: A whole-brain analysis revealed a network of regions including foci in the left medial prefrontal cortex (mPFC); right anterior cingulate cortex (ACC); and the precuneus. (Medial views of left and right hemispheres; for the full list of significant foci see Supplemental Table S2.) *Bottom panels:* Percent-signal-change averages according to subsequent memory and independent recognition performance during *SOL*. Green, activation while observing the solution of spontaneously identified images (*SPONT*); red, subsequently remembered (*REM*); blue, not remembered (*NotREM*).

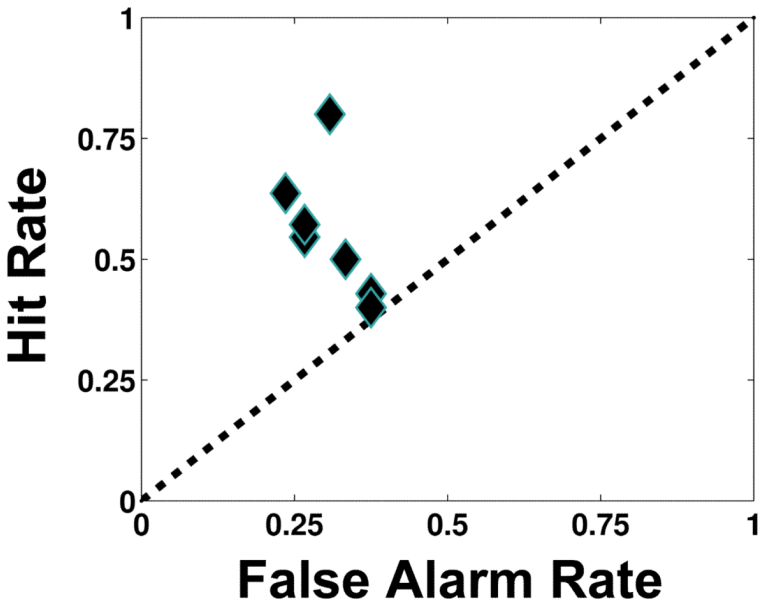


**Figure 8.**

Memory prediction accuracy. Hit rate (y-scale; i.e. the number of trials in which the image was predicted to be remembered and was indeed recognized at the Grid task one week later, as a fraction of the total number of *REM* trials) and false-alarm rate (x-scale; i.e. the number of trials in which the image was predicted to be remembered, yet was not recognized at the Grid task one week later, as a fraction of the total number of *NotREM* trials) relation per subject. Each [blue/black] circle represents a participant. The dotted line depicts chance-level accuracy. For the *'study'* protocol of *Experiment 3*, which was slightly different from that of *Experiment 2*, a view of the slice prescription used, and the time-courses of the amygdala ROIs see Supplemental Figure S3.