Adaptive Top–Down Suppression of Hippocampal Activity and the Purging of Intrusive Memories from Consciousness

Roland G. Benoit[1], Justin C. Hulbert[2], Ean Huddleston[3], and Michael C. Anderson[3]

Abstract

■ When reminded of unwanted memories, people often attempt to suppress these experiences from awareness. Prior work indicates that control processes mediated by the dorsolateral prefrontal cortex (DLPFC) modulate hippocampal activity during such retrieval suppression. It remains unknown whether this modulation plays a role in purging an intrusive memory from consciousness. Here, we combined fMRI and effective connectivity analyses with phenomenological reports to scrutinize a role for adaptive top–down suppression of hippocampal retrieval processes in terminating mnemonic awareness of intrusive memories. Participants either suppressed or recalled memories of pictures depicting faces or places. After each trial, they reported their success at regulating awareness of the memory. DLPFC activation was greatest when unwanted memories intruded into consciousness and needed to be purged, and this increased engagement predicted superior control of intrusive memories over time. However, hippocampal activity was decreased during the suppression of place memories only. Importantly, the inhibitory influence of the DLPFC on the hippocampus was linked to the ensuing reduction in intrusions of the suppressed memories. Individuals who exhibited negative top–down coupling during early suppression attempts experienced fewer involuntary memory intrusions later on. Over repeated suppressions, the DLPFC–hippocampus connectivity grew less negative with the degree that they no longer had to purge unwanted memories from awareness. These findings support a role of DLPFC in countermanding the unfolding recollection of an unwanted memory via the suppression of hippocampal processing, a mechanism that may contribute to adaptation in the aftermath of traumatic experiences. ■

INTRODUCTION

Not everything stored in memory is something that we wish was there. The unpleasantness that sometimes accompanies human experience is all too often remembered, and when we are reminded of these unwelcome events, we often strive to exclude them from awareness. Recent years have witnessed accumulating evidence that targeted efforts to suppress an experience from awareness can weaken the underlying memory representations of the excluded trace and eventually cause forgetting (Küpper, Benoit, Dalgleish, & Anderson, 2014; Noreen & MacLeod, 2013; Anderson & Huddleston, 2011; Anderson & Green, 2001). In this article, we examine the neural mechanisms supporting the purging of an unwanted memory at moments when that memory involuntarily enters awareness. We seek, in particular, to examine the mechanisms that tie the regulation of awareness to reductions in a memory's later propensity to intrude into consciousness. In so doing, we hope to provide a foundation for understanding how people adapt memory by attenuating involuntary recollections that often pervade their experience in the aftermath of unpleasant life events.

Much is already known about suppressing unwanted memories. For instance, attempts to suppress retrieval are associated with increased activation in right dorsolateral prefrontal cortex (DLPFC) and diminished activation of the hippocampus (HC; Benoit & Anderson, 2012; see also Gagnepain, Henson, & Anderson, 2014; Paz-Alonso, Bunge, Anderson, & Ghetti, 2013; Depue, Curran, & Banich, 2007; Anderson et al., 2004). Given the fundamental role of the latter region in recollection (e.g., Eichenbaum, Yonelinas, & Ranganath, 2007; Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000), these data suggest that the DLPFC exerts inhibitory control over critical retrieval processes supported by the HC. Effective connectivity analyses provided evidence for this direct suppression mechanism (Benoit & Anderson, 2012): The DLPFC influenced hippocampal activation, and a more negative coupling between these structures was associated with greater forgetting of unwanted memories (see also Gagnepain et al., 2014). Importantly, this negative modulation of HC activation did not arise for participants asked to avoid the unwanted memory by generating a distracting thought; it only arose for those told to directly suppress the retrieval

[1]Harvard University, Cambridge, MA, [2]Princeton University, NJ, [3]MRC Cognition and Brain Sciences Unit, Cambridge, UK

Journal of Cognitive Neuroscience 27:1, pp. 96–111
doi:10.1162/jocn_a_00696

process (for suppression vs. distraction, see also Bergström, de Fockert, & Richardson-Klavehn, 2009; see Anderson & Hanslmayr, 2014, for a review). Because negative modulation of HC activity was associated with instructions to terminate episodic retrieval, the inhibitory control mechanisms supporting this process may underpin the subjective experience of intentionally terminating awareness of an unwelcome reminding.

The current study examines the linkage between the engagement of top–down fronto-hippocampal modulatory processes and the regulation of awareness. Specifically, we test the possibility that this mechanism gets adaptively upregulated when intruding memories need to be purged from consciousness. Although existing data on retrieval suppression are consistent with this possibility, the key inference for this hypothesis is indirect. In Benoit and Anderson (2012), for example, whether memories intruded during suppression attempts is unknown and even if this was known, it remains possible that a different mechanism purged the memory from awareness, in parallel to the modulation of HC activation by DLPFC. HC modulation might instead serve a proactive function of preventing recollection. To establish that DLPFC is involved in purging momentary awareness (a form of adaptive control), requires that we link the occurrence of intrusions and their control to the DLPFC–HC suppression mechanism.

Suggestive evidence already exists for an adaptive control process that reactively purges memories from awareness via hippocampal downregulation. A recent study observed reduced HC activation when people suppressed unwanted memories but only when those memories intruded into awareness, as indexed by trial-by-trial intrusion judgments (Levy & Anderson, 2012). Unfortunately, the control process that reduced HC activation during intrusions was not strongly established in that study, leaving it uncertain whether the DLPFC mechanism was involved in purging memories from awareness.

In this study, we seek to identify two key features of the DLPFC–HC mechanism that would provide evidence for adaptive control. First, the top–down influence exerted by the DLPFC should be dynamically upregulated during intrusive memories, and such reactive engagement should contribute to reduced involuntary awareness of the intruding memory over time. Second, if this mechanism renders the memories less intrusive, there would be less need for inhibitory top–down control after repeated suppression attempts for a given memory. For instance, in research on retrieval-induced forgetting, PFC regions are recruited most during the initial retrieval of a target memory when competition from an interfering memory needs to be overcome (Wimber et al., 2011; Kuhl, Dudukovic, Kahn, & Wagner, 2007); critically, activation in those regions is attenuated during later retrieval attempts, when the competing memory had presumably been weakened and thus had become less interfering. On the basis of these findings, Kuhl et al. (2007) argued that the diminished need to engage cognitive control over repetitions was an adaptive benefit of forgetting competing memories. If so, an analogous benefit of successful retrieval suppression may be observed here, with diminished engagement of control over repeated suppression attempts, as intrusions of unwanted memories are downregulated. Thus, the DLPFC–HC suppression mechanism could be considered an adaptive control process if it gets dynamically upregulated to contend with involuntary intrusions, and it becomes progressively disengaged over trials with diminishing need to purge unwanted memories from awareness.

Participants first extensively encoded cue–picture associations via a procedure designed to induce strong and intrusive memories. Half of the pictures depicted faces; the other half, places. Using two types of material allowed us to examine whether the prefrontal inhibition mechanism is engaged irrespective of the mnemonic content (Benoit, Gilbert, Frith, & Burgess, 2012; Burgess et al., 2006). After training, participants performed a think/no-think task (Anderson & Green, 2001), as they were scanned with fMRI. They were presented with the cues from studied pairs and repeatedly recalled some associated pictures (recall condition) and attempted to suppress others (suppress condition). Critically, after each trial, participants indicated whether the picture had entered awareness (Levy & Anderson, 2012). This method of introspection, introduced in attention research (Corallo, Sackur, Dehaene, & Sigman, 2008; Sergent & Dehaene, 2004), enables us to identify activation associated with intrusions (i.e., picture involuntarily came to mind) versus nonintrusions (i.e., picture was kept out of awareness). It provides our central behavioral dependent measure, as it may be used to quantify the degree to which participants progressively gained control over intrusive memories.

If DLPFC supports an adaptive control mechanism that is instrumental to purging the contents of mnemonic awareness, we expected several core findings, both on individual trials as a function of intrusions and over the course of the think/no-think phase as a function of a diminishing intrusion frequency. First, DLPFC should be more robustly engaged when people need to countermand intruding memories than on trials during which they do not experience intrusions. If this mechanism is instrumental in purging memories from awareness, stronger engagement of DLPFC during intrusions may be associated with a greater decline in the memories' intrusiveness over time. Second, effective connectivity analyses should reveal a modulatory influence of DLPFC on HC during suppression. If a negative top–down coupling is implemented to purge unwanted memories from awareness, its engagement may be tied to the degree that such memories intrude into consciousness. It should be especially pronounced when individuals efficiently cope well with unwanted memories, as indicated by a greater subsequent reduction in intrusions. As a consequence of such efficient coping, there would subsequently be less demand for the purging mechanism, and we accordingly expect a successive attenuation of negative

top–down control over trials (cf. Wimber et al., 2011; Kuhl et al., 2007).

METHODS

Participants

All 18 volunteers were right-handed and not color blind, reported no history of neurological or psychiatric disorder, and gave written informed consent as approved by the local ethics research committee. Two participants were excluded: one because of poor recall of faces and one for noncompliance with the instructions as assessed by a postexperimental questionnaire (derived from Hertel & Calcaterra, 2005). Thus, data from 16 participants remained for further analyses (8 men, mean age = 22 years, range = 18–30 years).

Procedure

The procedure was based on the think/no-think task (Anderson & Green, 2001) and was divided into three phases (Figure 1A): (a) an elaborate study phase, during which participants encoded cue–memory associations; (b) the critical think/no-think phase, during which we assessed the key behavioral measure, that is, the experience of involuntary memory intrusions; and (c) a final recognition phase.

During study, participants encoded 24 critical cue–picture pairings. The cues were short nouns, and all pictures were presented in grayscale. Half of the images depicted unfamiliar faces (i.e., white female and male faces); the other half, places (water, tree, mountain, and prairie scenes). The pictures were assigned to the suppress, recall, and baseline conditions, counterbalanced across participants. Participants also encoded 12 filler pairings for practice. The study phase was composed of two stages designed to create strong memories that would likely intrude into awareness involuntarily. First, each cue–picture pairing was presented for 6 sec. Second, we ensured successful learning by a category recall task (i.e., presenting each cue word and asking participants to classify the retrieved scenes as a face or place) followed by a four-choice recognition task. The foils were sampled from

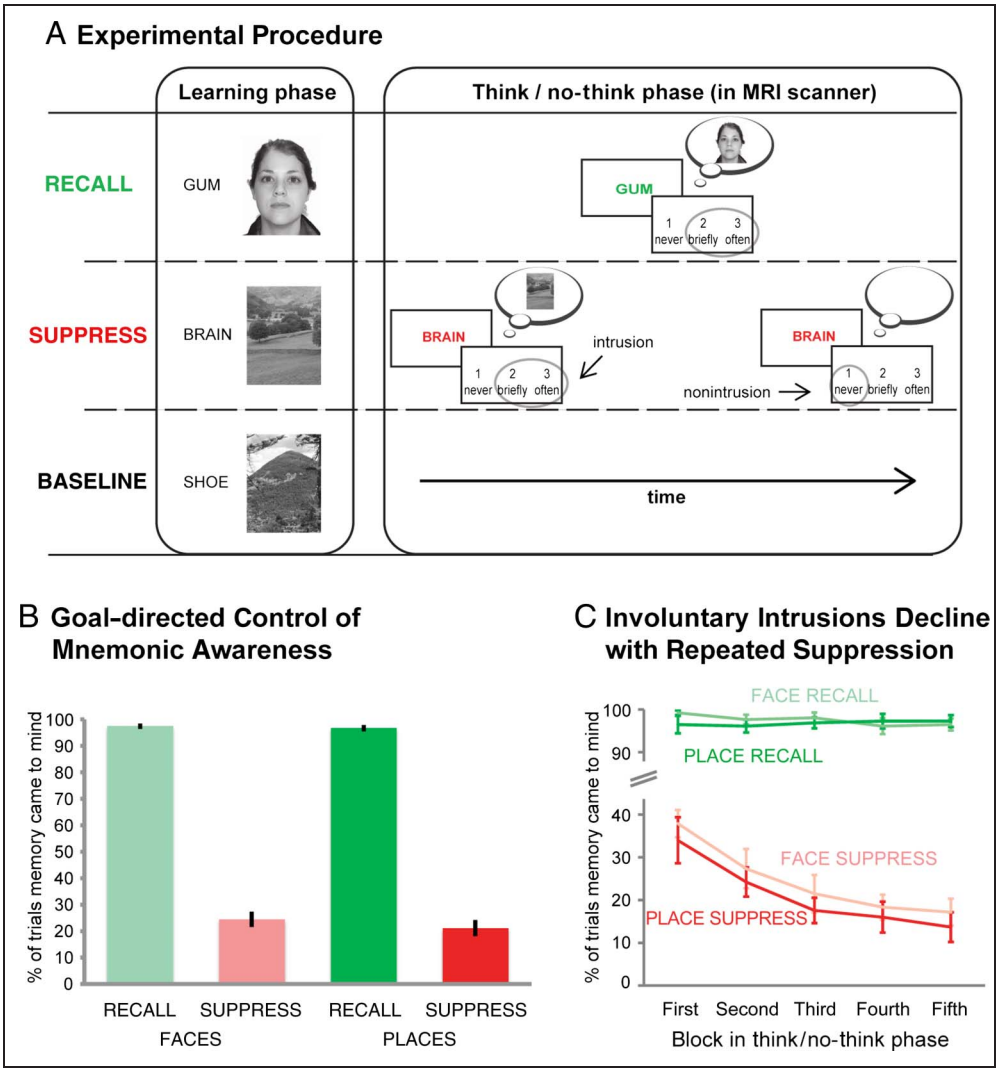

Figure 1. (A) Participants were trained on reminder–picture pairings, before they suppressed memories cued with red reminders and recalled those cued by green reminders. After each trial, they indicated whether the memory did come to mind (either briefly or often) or did not enter awareness at all. These critical trial-by-trial reports allowed us to identify brain activation associated with intrusions and to determine the degree to which participants gained better control over involuntary retrievals with repeated suppression attempts. (B) Whereas participants reported awareness of the reminded memory on most trials of the recall condition, they did experience intrusions of to-be-suppressed memories on a smaller number of trials. (C) The intrusion frequency of these items decreased over blocks of the experiment.

the pairings and were of the same type as the critical picture (e.g., male faces). By this, participants had to remember exact perceptual features. They received feedback after each test and were shown the cues along with the correct pictures for 2 sec at the end. A cue–picture pairing was removed from the training set, only if it had been recognized and its category had been recalled on the same trial. We thus ensured accurate memory for all pictures and further strengthened their representations by feedback.

After a practice session, participants performed the think/no-think phase in the MRI scanner. We presented cues in either red or green for 3 sec. Participants always read and comprehended the cue first. For green cues, they then covertly recalled the associated pictures (recall condition). By contrast, for red cues, they avoided thinking of the associated picture at all (suppress condition), in whatever way they saw fit (as in Anderson et al., 2004; Anderson & Green, 2001). Participants were thus free to adopt their own strategy, as long as they did not shift their eyes or attention away from the cue. After cue offset, they reported the extent to which they had thought about its associated picture by pressing one of three buttons: never, briefly, and often. This was followed by a jittered ISI ($M \pm SD = 3500 \pm 1588$ msec), optimized to increase the efficiency of the event-related design using optseq2 (surfer.nmr.mgh.harvard.edu/optseq/). There were five blocks, during each of which all recall and suppress cues were randomly presented twice, although any cue could only be repeated once all the others had been shown. Thus, in total, each suppress and recall memory was cued 10 times. We did not present cues for baseline memories during this phase.

Outside the scanner, after-effects of suppression on recognition memory were assessed. Given the extensive training of the memories and based on pilot data, we expected nearly perfect recognition accuracy. We therefore examined whether repeated purging from awareness slowed subsequent recognition. A trial started with a cue presented in white and the prompt to recall the category of the associated picture within 5 sec. The response was immediately followed by a blank screen (500 msec), before we presented the correct picture and a single foil. Participants had up to 5 sec to respond on this forced-choice recognition task. Analyses focused on correct recognition time, conditioned on items for which participants had also correctly recalled the category (cf. Detre, Natarajan, Gershman, & Norman, 2013).

fMRI Acquisition and Preprocessing

A 3-T Siemens TIM Trio MRI scanner was used for acquisition of T2*-weighted EPIs (64 × 64; 3 × 3 mm pixels; 3-mm thick, oriented to the AC–PC plane; repetition time = 2 sec; echo time = 30 msec; flip angle = 78°). Five runs were acquired, with the number of volumes varying from 193 to 226 because of the self-paced intrusion ratings. In addition, magnetization prepared rapid gradient echo structural images were acquired (256 × 240 × 192; 1-mm^3 isotropic voxels; repetition time = 2,250 msec; echo time = 2.99 msec; flip angle = 9°).

Preprocessing and univariate analyses were performed with SPM8 (www.fil.ion.ucl.ac.uk/spm/software/spm8). The volumes were realigned, corrected for different slice acquisition times, and coregistered with the structural image. This was spatially normalized, and the resulting parameters served to normalize the functional images into 3 × 3 × 3 mm^3 voxels by fourth degree B-spine interpolation (using the Montreal Neurological Institute reference brain). The images were then smoothed by an isotropic 8-mm FWHM Gaussian kernel.

fMRI Analyses

The variance in BOLD signal was decomposed in a general linear model (Friston et al., 1995) on the concatenated data from all runs to optimize estimation of intrusion events. Six regressors coded the 3 sec of each trial, separately for recall, nonintrusion (i.e., suppress trials with a "never" rating), and intrusion (i.e., other suppress trials) events of the faces and places trials. These regressors were convolved with the canonical hemodynamic response function. The full model was also composed of regressors for each run representing residual movement artifacts. We modeled trends over time by including, for each run, a linear-trend predictor, a six-predictor Fourier basis for nonlinear trends (sines and cosines of up to three cycles per run), and a confound-mean predictor (Kriegeskorte, Mur, & Bandettini, 2008). Parameters were estimated from the least mean-squares fit of the model to the data. For second-level analyses, contrast estimates for recall, nonintrusion, and intrusion events (for both faces and places) were entered into a repeated-measures ANOVA using nonsphericity correction, with Individuals as between-subject factor.

We modeled effective connectivity with DCM12 as implemented in SPM12b, following the procedure employed by Benoit and Anderson (2012). DCM explains regional effects in terms of changing patterns of connectivity during experimentally induced contextual modulations (Stephan et al., 2010). It allows for inferences about the presence and directions of causal connections, such as whether activity in the HC is influenced by activity in the DLPFC during suppression attempts (Benoit & Anderson, 2012). These inferences are based on evidence from Bayesian model selection (BMS; see below), which indicates the probability that a given model is more likely to have generated the data than the other models (e.g., a model that entails a modulation of the connectivity from DLPFC to HC vs. a model with a modulation of the reversed connectivity from HC to DLPFC; Penny et al., 2010). DCM requires univariate effects of interest in the respective brain areas (Stephan et al., 2010). Given that we did not observe hippocampal modulation during the suppression of face pictures (see Results), our analyses of effective connectivity are thus restricted to place data.

All models were variations of a standard model that is composed of the two regions (DLPFC and HC) as nodes, within-region inhibitory auto-connections, and bidirectional intrinsic connections (representing the connectivity between the regions across all conditions). Any cue to recall or to suppress could elicit responses in this network. This driving input was modeled as a series of delta functions at the respective cue onsets. It could enter the network either via the HC, the DLPFC, or both nodes, thus constituting three basic models (Figure 3A). Critically, we wanted to assess whether models could account better for the data that feature a modulation of the top–down connectivity from DLPFC to HC during attempts to suppress an unwanted memory. We thus created four model families, each of which was composed of modified versions of all three basic models. Across the model families, we varied the connection that could be modulated during suppression (Figure 3A). The modulatory component reflects the change in coupling from the average connectivity that arises specifically during suppression attempts. Family 1 did not have any such modulatory component, whereas Family 2 allowed for modulation of the connection from HC to DLPFC. Thus, these two families do not feature a modulation of the top–down connection, and they are incongruent with the hypothesized suppression mechanism. By contrast, Family 3 entailed modulation of the connection from DLPFC to HC, and Family 4 allowed both the DLPFC-to-HC and HC-to-DLPFC connections to be modulated during suppress events. Thus, these latter two families are congruent with the hypothesized suppression mechanism. The modulatory input was defined as changes in connectivity induced during the first second after the onset of a suppress event.

The models were estimated separately for each participant and for each of the five functional runs. By this, we were able to assess changes in top–down coupling across the think/no-think phase. We first identified the group peaks for the suppress-versus-recall contrast within a sphere (r = 10 mm) centered on the DLPFC peak previously associated with direct memory suppression (x, y, z = 36, 38, 31; Benoit & Anderson, 2012; suppress > recall) and within an anatomical mask of the right HC (Malian, Laurienti, Kraft, & Burdette, 2003; suppress < recall). Both peaks survived small-volume family-wise error correction (DLPFC: 39, 29, 34, z_{max} = 3.99; HC: 33, −25, −11, z_{max} = 4). We then identified the subject-specific peaks within a sphere (r = 10 mm) centered on the observed group peaks (Stephan et al., 2010). For the HC, the individual peaks also had to be within the anatomical mask. The individual peaks then served as centers for subject-specific, spherical ROIs (r = 5 mm). The first eigenvariate from an ROI (i.e., the first principal component of the time series of the voxels), adjusted for the effects of interest, constituted the regional activation. Model fitting was based on these data and was achieved by adjusting the parameters to maximize the free-energy estimate of the model evidence (Friston, Harrison, & Penny, 2003).

On the estimated models, we ran BMS in a random-effects approach, which does not assume that the optimal model is identical across all participants and which is also less susceptible to outliers than a fixed-effects approach (Stephan et al., 2010). This procedure reports the exceedance probability (EP), that is, the probability to which a given model is more likely than the others to have generated the data from a randomly selected participant (Penny et al., 2010). The EPs of all included models sum to 1. Note that BMS penalizes for model complexity. We then analyzed the effective connectivity (i.e., the sum of the intrinsic and modulatory parameters) of the winning model family, to examine the relationship between top–down coupling during suppression and control of involuntary intrusions. Specifically, we performed Bayesian model averaging (BMA; Penny et al., 2010), which computes weighted averages of each model parameter, where the weighting is determined by the posterior probability of each model. To examine changes in connectivity across the think/no-think phase, we conducted BMA separately for each of the five blocks. We also performed BMA on models from the first and second blocks (approximating the first half of the think/no-think phase) to assess the relationship between negative coupling during this period and subsequent reductions in involuntary intrusions.

RESULTS

Behavioral Results

Repeated Suppression Attempts Decrease the Intrusion of Unwanted Memories

We first analyzed the frequency with which people reported awareness of the associated memory on both recall and suppress trials. As might be expected based on the extensive training of pairs, on recall trials, participants reported they were nearly always successful at bringing the associated picture into awareness during the think/no-think phase ($M \pm SD$: faces = 98 ± 4%; places = 97 ± 5%). Moreover, the frequency of reported retrieval did not change reliably across the session (i.e., from the first block to the last block) as indicated by an ANOVA with the factors Block and Material (Block: $F(1, 15) = 1.44$, $p = .25$).

On suppress trials, participants reported that pictures came into awareness far less frequently (faces: 24 ± 12%; places: 21 ± 12%) than during recall trials (Figure 1B). This dramatic difference in reported mnemonic awareness was corroborated by an ANOVA with the factors Retrieval goal (recall, suppress) and Material (faces, places) that yielded a significant effect of Retrieval Goal only ($F(1, 15) = 699.31$, $p < .001$). Thus, participants exhibited a highly robust ability to control awareness of the pictures (a nearly 75% difference in awareness across conditions). Nevertheless, the images intruded frequently during suppression trials, although the intrusion frequency declined from the first block to the last block (cf. Levy & Anderson, 2012): An ANOVA with the factors Block and Material

yielded a substantial effect of Block ($F(1, 15) = 60.3$, $p < .001$), reflecting diminished intrusion frequencies for both faces and places with repeated suppression attempts (Figure 1B). Neither the Material effect nor the interaction was significant (all Fs < 1.09, all ps $> .3$). Thus, our participants showed evidence of substantial improvements in coping with intrusive memories, irrespective of mnemonic content. This pattern may reflect a progressive weakening of the memories because of repeated suppression. Our central objective is to understand the mechanisms underlying this highly robust reduction in intrusive memories over time and whether an adaptive upregulation of the modulatory DLPFC influence on HC contributes to it.

Suppressing Unwanted Memories Slows Later Recognition

We expected nearly perfect performance on our final test because of our extensive training procedure. Indeed, people's performance at classifying the category of the picture in response to the cue word (e.g., as either a face or a place) was very high ($M \pm SEM$: 0.92 ± 0.02 for places, 0.94 ± 0.02 for faces), and an ANOVA with the factors Retrieval goal (suppress, recall, baseline) and Materials (faces, places) yielded no reliable differences or interactions (all Fs < 1.3, all ps $> .28$). Similarly, as expected based on the extensive training and piloting results, recognition performance was virtually perfect across all conditions (Table 1; all Fs $= 1.06$, all ps $= .319$).

On measures of recognition time, however, an ANOVA with the factors Retrieval goal and Materials revealed a significant effect of Retrieval goal on response times ($F(2, 30) = 7.42$, $p < .003$; Table 1), and a follow-up ANOVA indicated that responses for suppress items were slower than those for baseline items ($F(1, 15) = 8.78$, $p < .02$). Thus, people were slower to recognize suppressed pictures. Surprisingly, however, we also found that recall items were recognized more slowly than baseline items ($F(1, 15) = 16.31$, $p < .002$), which could not have been produced by a suppression process. The main effects of Material were also significant, reflecting faster responses for faces (all Fs > 8.96, all ps $< .01$). Slowed RTs for suppress and recall items (compared with baseline items) did not correlate (faces: $r(14) = .098$, $p = .719$; places: $r(14) = -.003$, $p = .99$), suggesting that the slowed recognition for these two item types may reflect different mechanisms (see Discussion). However, exploratory analyses did not yield any significant relationships between suppression-induced slowing and any of the brain measures linked to the efficient coping with unwanted memory intrusions reported in the next sections.

Table 1. Recognition Performance

	Recognition		*Recognition Times (msec)*	
	Mean	*SEM*	*Mean*	*SEM*
Faces				
Baseline	0.97	0.01	1219	90
Recall	0.95	0.02	1384	93
Suppress	0.95	0.03	1341	103
Places				
Baseline	0.96	0.03	1452	81
Recall	0.95	0.02	1652	110
Suppress	0.93	0.03	1610	111

fMRI Results

DLPFC Activation Is Greater during Memory Intrusions

To establish the involvement of a content-general suppression mechanism in DLPFC, we performed a conjunction analysis of the suppress-versus-recall contrasts. This analysis averaged estimates of intrusion and nonintrusion trials and thus reveals those regions that are engaged across both types of suppression trials. We then performed a conjunction analysis of the contrasts for the faces and places items to identify regions showing this effect for both material types. Critically, an extensive right DLPFC area centered on the middle frontal gyrus survived whole-brain cluster correction (peak: $x, y, z = 42, 26, 37$; $z_{max} = 4.23$; 160 voxels; Figure 2A). This area included those voxels previously implicated in direct memory suppression (i.e., around $x, y, z = 36, 38, 31$; Benoit & Anderson, 2012; for complete whole-brain analyses, see Tables 2–5). Thus, a common DLPFC region is robustly engaged during the suppression of images of faces and places and words (as previously shown).

Having identified this control region, we next examined whether it functioned to proactively gate retrieval or, instead, could be adaptively upregulated to purge intruding mental content from awareness. The previous analysis collapsed across contrast estimates for all suppression trials, that is, irrespective of the occurrence of intrusions. The obtained DLPFC cluster is thus derived from a contrast vector that is orthogonal to the comparison of intrusion versus nonintrusion trials. Hence, we extracted contrast estimates for intrusions and nonintrusions from this cluster and conducted an ANOVA with the factors Awareness (intrusion, nonintrusion) and Material (faces, places). This analysis revealed greater DLPFC activation when participants experienced intrusions than when they did not ($F(1, 15) = 7.68$, $p < .02$; Figure 2B). Thus, although DLPFC was generally engaged when participants suppressed memories, it was more strongly recruited in situations that required the countermanding of intruding memories. However, there was a trend for the interaction between this effect and material type ($F(1, 15) = 3.53$, $p = .08$).

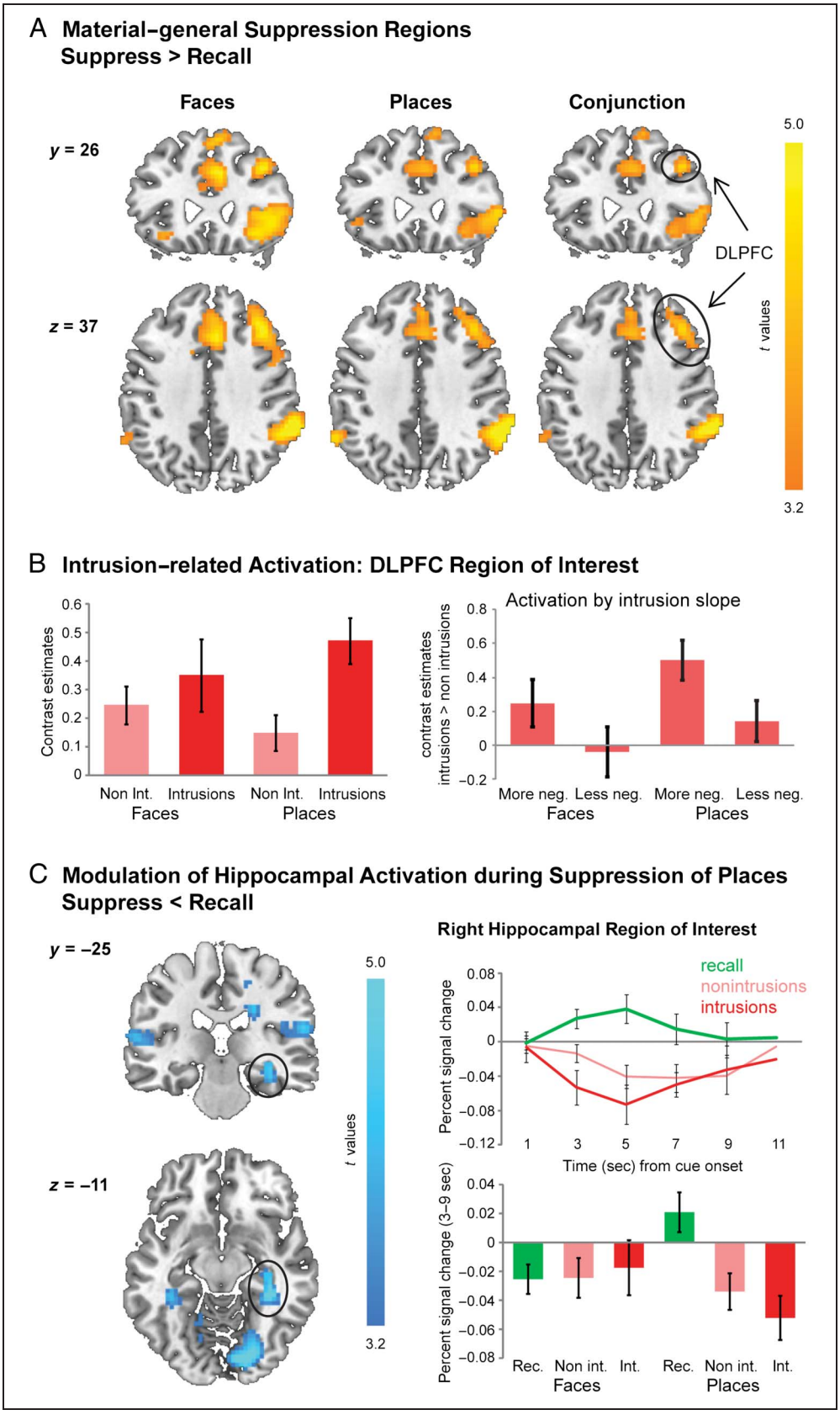

Figure 2. (A) Collapsing across estimates for intrusion and nonintrusion trials, DLPFC showed stronger activation during attempts to suppress (vs. recall) memories of both faces and places (conjunction analysis). (B) Consistent with a reactive account of memory suppression, activation in the DLPFC cluster was greater during intrusions than nonintrusions (left), and the activation difference between intrusions and nonintrusions was greater for individuals who were more successful at decreasing intrusion frequency over time (i.e., with a more negative intrusion slope; right). (C) The right HC yielded the expected reduced activation during suppression of place memories only. For display purposes, statistical maps are thresholded at $p < .001$, uncorrected.

Table 2. Regions in Which BOLD Signal Was Significantly Greater during Suppression than Recall of Place Memories

			MNI (Peak)				
Region	*~BA*	*Hemi.*	*x*	*y*	*z*	*Voxels*	Z_{max}
SFG/MiFG	10/11	l	−24	59	−8	14	3.69
MiFG/SFG	10	r	27	56	22	75	4.09
ACC	32	r/l	3	35	31	172	4.21
			3	23	37	Same cluster	3.91
MiFG/SFG	9	r	42	26	37	195	4.23
			30	38	37	Same cluster	3.65
			51	14	40	Same cluster	3.64
IFG	45/47	r	57	26	4	273	4.22
			33	23	−14	Same cluster	4.05
STG	13	r	42	17	−11	Same cluster	4.01
SFG/MeFG	6/8	r	12	23	61	120	4.19
			18	11	70	Same cluster	3.45
			6	11	58	Same cluster	3.45
IFG	47	l	−45	23	−5	48	3.76
			−33	23	−14	Same cluster	3.52
PreG/PosG	4/3	l	−51	−13	46	31	3.98
IPC	40	r	66	−43	40	268	5.31
SMG/IPC	40	l	−60	−52	37	34	4.24

Threshold at $p < .001$, uncorrected, and at least 10 contiguous voxels. ACC = anterior cingulate cortex; IFG = inferior frontal gyrus; IPC = inferior parietal cortex; MeFG = medial frontal gyrus; MiFG = middle frontal gyrus; PosG = postcentral gyrus; PreG = precentral gyrus; SFG = superior frontal gyrus; SMG = supramarginal gyrus; STG = superior temporal gyrus; Hemi = hemisphere; r = right; l = left.

Stronger DLPFC Upregulation Predicts Better Coping with Intrusive Memories

Purging a memory from awareness may be achieved by inhibitory processes that reduce its tendency to intrude again on later trials. If so, participants who recruit the DLPFC region more during intrusions than during nonintrusions may exhibit a steeper decline in intrusion frequency over repeated suppression attempts. To test this prediction, we performed a median split of our sample based on the baseline-corrected intrusion slope (i.e., the decrease in intrusion frequency over repetitions of suppression, divided by the initial frequency of intrusions, as reported by Levy & Anderson, 2012). By correcting for individual differences in the initial intrusion frequency, this measure allows for the definition of groups who were high or low in their ability to cope with intrusive memories over time. Given the absence of interactions between material type and the decline in intrusion frequency (see above), this split was performed on intrusion slopes averaged across places and faces data. The individuals exhibiting a more negative slope trended to experience intrusions on fewer trials overall ($M \pm SEM = 18 \pm 3\%$ vs. $28 \pm 4\%$; $t(14) = -1.97$, $p = .069$), consistent with a superior control ability.

We then performed an ANOVA on the measure of DLPFC upregulation (i.e., the contrast estimates for intrusions > nonintrusions) with the between-subject factor Intrusion slope (more negative, less negative) and the within-subject factor Material (places, faces). Consistent with the prediction, participants who more effectively coped with intrusions, reducing their frequency over time (i.e., a more negative slope), showed a significantly greater intrusion-related upregulation of DLPFC activation ($F(1, 15) = 5.83$, $p < .05$; Figure 2C). This effect did not interact with Material type ($F(1, 15) = 0.1$, $p = .76$). Thus, a greater increase in DLPFC activation when memories intruded into awareness was associated with a greater decline in intrusion frequency, irrespective of the nature of the unwanted memories.

Hippocampal Activation Is Reduced during Attempts to Suppress Unwanted Place Memories

The hypothesized purging mechanism is thought to suppress awareness of intruding memories via the modulation

of HC activation by DLPFC. Because DLPFC has been found to modulate right HC during direct suppression (Benoit & Anderson, 2012), we extracted contrast estimates for recall, nonintrusion, and intrusion trials from an anatomical mask of this region (Malian et al., 2003). In the current data, hippocampal activation seemed to be reduced during attempts to suppress place memories only (Figure 2C). This impression was corroborated by an ANOVA with the factors Retrieval status (recall, nonintrusions, intrusions) and Material (faces, places) that yielded a significant interaction ($F(2, 30) = 4.15$, $p < .03$). For places, activation was lower during attempts to suppress (compared with recall) unwanted memories both in the absence ($t(15) = 4.24$, $p < .001$, one-tailed) and presence

Table 3. Regions in Which BOLD Signal Was Significantly Greater during Recall than Suppression of Place Memories

			MNI (Peak)				
Region	*~BA*	*Hemi.*	*x*	*y*	*z*	*Voxels*	Z_{max}
ACC	24	l	−18	32	1	11	3.63
ACC	23/32	r/l	6	23	−2	29	3.73
			−6	20	−2	Same cluster	3.49
STG/PosG	22/42/40	l	−48	−10	7	184	3.94
			−63	−25	13	Same cluster	3.92
			−57	−22	7	Same cluster	3.63
PHC		r	39	−13	−26	18	3.6
			30	−13	−17	Same cluster	3.25
Striatum		l	−21	−19	31	10	3.34
			−21	−7	28	Same cluster	3.32
Striatum		r	24	−25	31	41	4.01
PosG	40	r	60	−28	19	312	4.65
Insula	13	r	48	−7	13	Same cluster	4.1
STG	41	r	57	−19	10	Same cluster	4.03
CG	31		6	−34	34	10	3.67
Insula	13	l	−27	−34	16	119	4.3
PHC	37	l	−36	−40	−8	Same cluster	4.24
HC			−27	−37	4	Same cluster	3.26
PHC	36/37	r	36	−34	−14	237	5.19
Striatum			30	−34	10	Same cluster	4.51
			36	−43	4	Same cluster	4.03
HC			30	−31	−2	Same cluster	3.48
CG	31	l	−12	−37	46	43	3.82
Cerebellum			−24	−37	−47	25	3.63
MTG/SOG	39/19	r	42	−76	22	45	4.03
MTG/SOG	39/19	l	−39	−82	22	67	4.2
			−33	−76	31	Same cluster	3.62
Cuneus	17/18	r	18	−91	4	3074	6.84
			15	−76	16	Same cluster	5.43
PCC	31	r	12	−67	13	Same cluster	5.29

Threshold at $p < .001$, uncorrected, and at least 10 contiguous voxels. CG = cingulate gyrus; PCC = posterior cingulate cortex; PHC = parahippocampal cortex; SOG = superior occipital gyrus. See Table 2 for additional acronyms.

Table 4. Regions in Which BOLD Signal Was Significantly Greater during Suppression than Recall of Face Memories

			MNI (Peak)				
Region	*~BA*	*Hemi.*	*x*	*y*	*z*	*Voxels*	Z_{max}
MiFG/SFG	10	r	33	53	22	1345	5.36
	6		15	14	61	Same cluster	5.08
ACC	32	r/l	6	32	31	Same cluster	4.81
IFG/insula	47/13	l	−30	20	−11	101	4.48
			−36	14	1	Same cluster	3.47
IFG/insula	45/47/13	r	48	17	4	533	5.34
			33	20	−11	Same cluster	4.99
			45	23	−5	Same cluster	4.49
Striatum		r	15	8	7	22	3.54
ITG	20	r	48	−1	−35	13	3.6
PreG/PosG	4/3	L	−48	−16	46	15	3.67
IPC/SMG	40	r	60	−46	25	391	5.21
			60	−43	34	Same cluster	5.05
			51	−49	34	Same cluster	4.68
Cerebellum		l	−33	−49	−32	37	3.84
SMG	40	l	−66	−52	31	18	3.74
IOG/MOG/ITG	19/18	l	−45	−79	−5	55	3.59
MTG/ITG	37	l	−45	−67	1	Same cluster	3.43
IOG	18/19	l	−36	−85	−8	Same cluster	3.33
MOG	19	l	−36	−94	13	38	3.89

Threshold at $p < .001$, uncorrected, and at least 10 contiguous voxels. IOG = inferior occipital gyrus; ITG = inferior temporal gyrus; MOG = middle occipital gyrus. See previous tables for additional acronyms.

of intrusions ($t(15) = 2.04$, $p < .03$, one-tailed), although inspection of the time course data (as in Levy & Anderson, 2012) indicates that intrusion trials were associated with numerically, yet not significantly, lower activation. In contrast, for faces, activation was reduced neither during nonintrusions ($t(15) = 0.91$, $p = .19$, one-tailed) nor intrusions ($t(15) = 0.03$, $p = .49$, one-tailed). Thus, whereas the right DLPFC was recruited during attempts to suppress both faces and places, we observed evidence for the complementary reduction in HC activation only during the suppression of places.

The Top–Down Connectivity from DLPFC to HC Is Modulated during Memory Suppression

During the suppression of place memories, we observed the activation pattern previously associated with direct

Table 5. Regions in Which BOLD Signal Was Significantly Greater during Recall than Suppression of Face Memories

			MNI (Peak)				
Region	*~BA*	*Hemi.*	*x*	*y*	*z*	*Voxels*	Z_{max}
ACC	25/24/32	l/r	−6	20	−2	18	3.73
			6	17	−5		3.24
SPC	7	l	−39	−76	46	24	4.05
Cuneus/MOG/LG	17/18	r	18	−94	4	132	4.6

Threshold at $p < .001$, uncorrected, and at least 10 contiguous voxels. LG = lingual gyrus; SPC = superior parietal cortex. See previous tables for additional acronyms.

memory suppression: increased DLPFC and reduced HC activation. We hypothesized that the DLPFC would cause this reduction to effectively purge unwanted memories from awareness and that such suppression would be associated with evidence of people's ability to efficiently cope with memory intrusions over time. Efficient coping would be reflected in two ways. First, we predicted a relationship between the effective top–down connectivity and changes in the frequency of involuntary intrusions. Those individuals who exhibit an inhibitory (i.e., negative) influence during early suppression trials should be more likely to disrupt intruding traces and, as a result, experience fewer involuntary intrusions on later suppression trials. Second, to the extent that memories are rendered less intrusive over time by early and efficient control, the need to engage top–down control may be reduced. Hence, we used DCM to examine the effective connectivity between the DLPFC and the HC during the suppression of place memories and how this changed over blocks. It was not possible to model DLPFC–HC interactions during face suppression because DCM requires all univariate effects of interest in the respective ROIs (Stephan et al., 2010).

DCM explains regional effects in terms of dynamically changing patterns of connectivity during experimentally induced contextual changes. Thereby, it allows inferences about the direction of causal connections, that is, whether suppress events are more likely to modulate the top–down connection from DLPFC to HC or the reverse, bottom–up connection. We first tested whether models could account best for the fMRI data that include a modulation of the top–down connectivity from DLPFC to HC during memory suppression, similar to what has been observed in prior work (Gagnepain et al., 2014; Benoit & Anderson, 2012). We then examined the consequences of this putative top–down influence. Specifically, we examined whether it could be linked to the reduction in involuntary intrusions with repeated suppression attempts.

As detailed in Methods, we created four model families that varied the connection that could be modulated during suppression (Figure 3A). Family 1 did not have any modulatory component, whereas Family 2 entailed a modulation of the connectivity from HC to DLPFC. Thus, these two families are incongruent with the hypothesized suppression mechanism because they do not feature a modulation of the top–down connection. By contrast, the remaining two families are congruent with the hypothesized increased top–down influence: Family 3 entailed a modulation of the connection from DLPFC to HC, and Family 4 allowed both the DLPFC-to-HC and HC-to-DLPFC connections to be modulated.

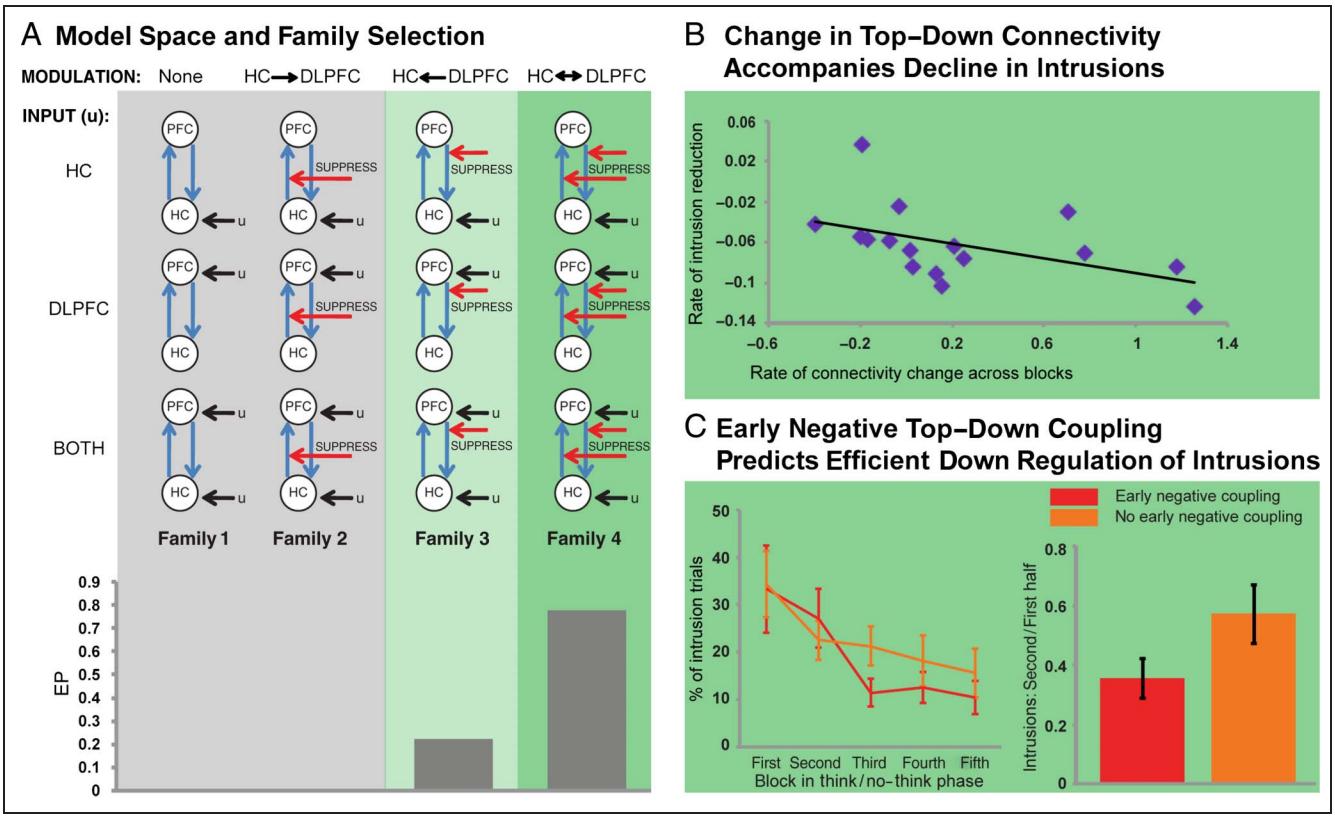

Figure 3. (A) BMS indicated that Family 4 could account best for the places data, consistent with the hypothesized increased impact of DLPFC on activation changes in the HC during suppression. (B) Changes in effective connectivity across the think/no-think phase were stronger for those individuals who exhibited a steeper decline in the number of involuntary intrusions, suggesting that a reduced requirement to purge unwanted memories leads to a disengagement of top–down inhibitory control. (C) Individuals exhibiting negative DLPFC–HC coupling early on (i.e., across the first two blocks) subsequently experienced a greater reduction of involuntary intrusions.

On the estimated models, we ran BMS, which returns EPs, that is, the probability for each model family to be more likely than the other families to have generated the data (Penny et al., 2010). The selection procedure followed the approach adopted by Benoit and Anderson (2012). First, we compared the four families, and BMS indicated that Family 4 was superior with an EP of 0.78 (Figure 3A). Thus, replicating Benoit and Anderson (2012), those models accounted best for the data that featured a modulation of the critical connection from DLPFC to HC during suppression as well as of the reverse, bottom–up connection. We then compared the three models within the winning Family 4, that is, the models that varied in the location of the driving input. Within this set, there was a clear preference for the model that received its driving input both via the HC and the DLPFC (EP > 0.99). To summarize, consistent with our hypothesis, the winning model featured a modulation of the connectivity from DLPFC to HC, suggesting that DLPFC influences HC activation specifically during memory suppression.

In the following, we further tested the adaptive control account by examining the nature of the effective top–down connectivity. The effective connectivity is the sum of the intrinsic (i.e., average connectivity) and modulatory connectivity parameters (i.e., change in connectivity specifically during suppress trials) and thus indicates the implementation of top–down control that is specific to memory suppression. BMA of the members of family IV indicated that, across the entire course of the think/no-think phase, the effective connectivity from the DLPFC to the HC was not reliably different from 0 ($M = 0.9$, $SEM = 1.15$; $t(15) = 0.78$, $p = .45$). However, as described next, this overall result masks a critical dynamic adjustment in top–down suppression over blocks predicted by the adaptive control hypothesis.

Top–Down Suppression Changes with Diminishing Need to Purge Involuntary Intrusions

People who cope efficiently with intrusions should show a rapid decline in intrusions across the think/no-think phase, reflecting the accumulating effects of control on the suppressed memory representations. As a result of mounting effects of control, the need to further engage top–down inhibitory control should decline as well, reflecting the comparative rarity of intrusion events that require action. The adaptive control account posits that this reduction in the need for control should be reflected in dynamic adaptation of DLPFC–HC coupling over blocks (cf. Wimber et al., 2011; Kuhl et al., 2007). Those individuals who experience a greater reduction in involuntary intrusions should progressively disengage this mechanism, as there is diminishing need to purge unwanted memories from awareness. Accordingly, changes in DLPFC–HC coupling over the course of the suppression phase should be tied to the rate that the intrusion frequency declines across the think/no-think phase.

We thus performed BMA of the members of Family 4 separately for each of the five runs and then calculated the slope of the top–down effective connectivity across blocks. This slope measure is positive for individuals who show a weaker negative coupling over blocks. Notably, some participants exhibited negative slopes, reflecting that they only developed a negative top–down connectivity with repeated suppression attempts.

Next, we examined the critical hypothesis that individuals who cope better with intrusions, and who therefore show a steep decline in intrusions over blocks (i.e., a more negative baseline-corrected behavioral intrusion slope), will show a decreasing need to implement top–down inhibitory control as blocks progress. The coupling between DLPFC and HC for efficient individuals should initially be negative and becomes less negative over blocks; thus, their change in coupling should be more positive. Taken together, the adaptive control hypothesis therefore predicts a negative correlation between behavioral intrusion slopes (which are more negative for highly efficient individuals) and coupling parameter slopes (which should be positive for efficient individuals who bring intrusions under better control).

We observed this predicted negative correlation between the decline in intrusions and change in connectivity (Figure 3B; $r(16) = -0.49$, $p < .03$, one-tailed). This suggests that participants who showed the most precipitous reduction in involuntary intrusions also showed a stronger disengagement of the suppression mechanism over blocks. Conversely, those who were least successful in reducing memory intrusions may have modulated HC less overall.

Large Reductions in Intrusions Are Predicted by Negative DLPFC–HC Coupling during Early Suppressions

The relationship between change in connectivity strength and decline in intrusion frequency suggests that an initially negative coupling leads to a reduction in the number of experienced intrusions on subsequent trials. Accordingly, individuals who already exhibit a negative top–down connectivity during earlier suppression attempts should experience fewer involuntary intrusions of the suppressed memories later on. Thus, the nature of the top–down connectivity during the first half of the think/no-think phase should predict the reduction in intrusions that participants experience during the second half. To test this prediction, we calculated the ratio of involuntary memory intrusions in the second versus first half, separately for each participant. This value is smaller for those who experience relatively fewer intrusions over time. We then split our sample based on the effective top–down connectivity parameter estimates to compare the intrusion reduction for those individuals who exhibited negative coupling parameters across the first two blocks of the think/no-think phase (i.e., approximating the first half of the session) and those who did not show evidence yet of implementing

an inhibitory suppression mechanism during that same period. Consistent with a role of inhibitory modulation during earlier trials in reducing involuntary intrusions on later trials, early negative couplers experienced a greater reduction in intrusions than individuals who did not exhibit early negative coupling (Figure 3C; $t(14) = -1.79, p < .05$, one-tailed).

Taken together, the data provide evidence for the adaptive deployment of the inhibitory top–down mechanism in the purging of unwanted memories from awareness. Activation in DLPFC was strongest when unwanted memories intruded into consciousness, and a greater upregulation of this activation was associated with a steeper decline of involuntary retrievals. Moreover, for place memories, a negative DLPFC-to-HC coupling during earlier suppression attempts predicted a greater reduction of involuntary intrusions on later trials. The coupling became weaker as there was diminishing need for a reactive control mechanism, further indicating that it was implemented adaptively when there was a greater demand to counteract involuntary retrievals.

DISCUSSION

When unwanted memories intrude into consciousness, people often seek to cope with these experiences by limiting the memories' duration in awareness. What mechanism achieves this purging, and how is it implemented in the brain? Past work provided evidence for a direct suppression mechanism that stops episodic retrieval, via negative coupling from DLPFC to HC (Gagnepain et al., 2014; Benoit & Anderson, 2012). Here, we showed that the impact of DLPFC on hippocampal activity can be upregulated adaptively to purge unwanted memories from awareness and that such counterintentional retrievals may actually foster the engagement of processes that minimize involuntary memory intrusions over time.

To isolate the role of the DLPFC in purging memories from awareness, participants reported memory intrusions on a trial-by-trial basis. Whereas all suppression attempts engaged the DLPFC, intrusions did so more robustly than did nonintrusions. This indicates that processes supported by DLPFC can be upregulated to countermand an unfolding recollection. This elevated intrusion response occurred independently of the mnemonic content (faces or places), suggesting that DLPFC supports a central process during memory suppression (cf. Burgess et al., 2006). Critically, the DLPFC engagement during intrusions appeared effective in reducing the later intrusiveness of the suppressed memories, predicting the decline in intrusions over suppression repetitions. This pattern may indicate that the reactivation of an intruding memory renders it more vulnerable to disruption (Detre et al., 2013; Levy & Anderson, 2012; Norman, Newman, & Detre, 2007), echoing research on reconsolidation (Dudai, 2004; Nader, Schafe, & Le Doux, 2000). In this reactivated state, inhibitory processes supported by DLPFC may be particularly effective in disrupting the memory's trace, for example, by distorting its hippocampal replay. This, in turn, would make it less likely for the suppressed memory to be reactivated involuntarily on later trials, as indexed by a decrease in memory intrusions.

The role of the DLPFC in reactive control received further support from connectivity analyses. The models that could account best for the place data were those that featured a modulation of the top–down connection from DLPFC to HC, consistent with prior work on direct suppression (Gagnepain et al., 2014; Benoit & Anderson, 2012). The effective connectivity between these regions changed with the rate that participants had gained better control over intruding memories. Those who had already exhibited a negative DLPFC–HC coupling during initial suppression attempts experienced less involuntary intrusions later on, and this top–down influence became progressively weaker as the demand to purge involuntary intrusions diminished. This pattern thus provides direct evidence for the central assertion of the adaptive control hypothesis (Kuhl et al., 2007), that is, that a benefit of weakening unwanted memories is the subsequent reduced demand on mnemonic control. Whereas previous support for this account was based on the assumption that successive reductions in prefrontal engagement were a consequence of successive weakening of the memories (Wimber et al., 2011; Kuhl et al., 2007), our online assessment of intrusions allowed us to directly link the rate of decreasing control with the rate of reduction in the memories' intrusiveness.

Together, these findings indicate that, for places at least, the DLPFC adaptively modulates HC activity particularly during intrusions and that a negative coupling contributes to diminished intrusiveness of suppressed memories. This modulation need not be direct, however, and may be achieved via relay nodes such as other medial temporal lobe structures or the retrosplenial cortex (Morris, Pandya, & Petrides, 1999; Goldman-Rakic, Selemon, & Schwartz, 1984), especially given the lack of evidence for monosynaptic connections from DLPFC to HC.

The best models in the current experiment also featured a modulation of the connectivity from the HC to the DLPFC during suppression, consistent with previous work (Benoit & Anderson, 2012). Given that activation in the HC may signal the involuntary retrieval of an unwanted memory (cf. Levy & Anderson, 2012), the HC-to-DLPFC connection could constitute a feedback loop that transmits the requirement for an upregulation of the inhibitory processes supported by the DLPFC.

Despite similar DLPFC activation for places and faces, we did not observe reliable reductions in HC activation when participants suppressed faces. Because of the absence of this effect, we could not use DCM for this class of stimuli. A recent study reported a similar pattern of material-specific differences during memory suppression (Detre et al., 2013): whereas a classifier of fMRI data yielded evidence for involuntary reactivation of place information, this was not the case for faces. More specifically, the

present lack of a difference in HC activation may reflect less dependence on HC processes during the retrieval of faces, as suggested by striking material-specific effects in amnesia showing that memory for unfamiliar faces may be supported by structures outside the HC (Aly, Knight, & Yonelinas, 2010; Bird & Burgess, 2008). The present lack of HC activation for faces on recall trials, despite participants reporting face retrieval on 97% of the trials, may be another instance of this material effect. As the neural basis of covert episodic recall of faces is not widely studied, this possibility must remain speculative. Our DCM evidence for HC modulation during the suppression of places converges, however, with similar modulations observed during the suppression of verbal materials (Benoit & Anderson, 2012) and visual objects (Gagnepain et al., 2014), indicating the generalization of this suppression mechanism across other material types.

The progressive weakening of inhibited memories, as indexed by a reduction in involuntary memory intrusions, may eventually also lead to impaired voluntary access. In the current study, we employed a training regimen that was designed to create strong memories overall; yet, on the final test, participants were slower to recognize pictures that they had previously suppressed compared with initially equally strong baseline memories. This finding adds to recent evidence that suppressing a memory can impede subsequent recognition (Kim, Oh, Kim, Sim, & Lee, 2013; Hart & Schooler, 2012; Waldhauser, Lindgren, & Johansson, 2012), thus extending previous evidence for suppression-induced deficits in recall (Anderson & Huddleston, 2011). Surprisingly, however, participants were also slower to recognize pictures that they had repeatedly recalled during the think/no-think phase, although they had reported recalling those memories on nearly all trials. Importantly, the slowed recognition of recall and suppress items did not correlate across subjects, suggesting a different underlying mechanism. Our extensive training of the cue–picture pairs may have contributed to this effect, because further attempts to retrieve Recall images during the Think/No-Think phase would have led to minimal further strengthening. Instead, additional retrieval may have contaminated the memory representations of the items by incorporating additional incorrect features (Schacter, Guerin, & St Jacques, 2011). Thus, on the final test, a mismatch between the probe image and its memory representation may have slowed recognition. Future research is required to understand the cause of this intriguing impairment.

Although our data indicate a role of DLPFC in purging a memory from awareness, DLPFC may also serve a proactive gating function that suppresses retrieval before an item reaches awareness. Several observations are consistent with this possibility. First, the frequency of intrusions on suppress trials was far lower than that of successful recalls on recall trials, even on the first repetition, suggesting a considerable ability to prevent retrieval. Second, although DLPFC was significantly more engaged by intrusions, it was also recruited in the absence of intrusions (compared with the recall condition), suggesting that it may not require awareness of the unwanted memory to engage this control process. Thus, whereas involuntary retrievals may trigger elevated control, control may be present in all cases.

This raises the possibility that a memory intrusion may simply be a proxy for a high level of reactivation of the unwanted memory. If so, enhanced direct suppression may not be triggered by an intrusion into awareness but rather may be engaged to the degree to which a memory gets reactivated, irrespective of awareness (Levy & Anderson, 2012). If suppression depends on a reactivation level, this may explain why HC activation was also reduced for nonintrusions, whereas it was not in prior work (Levy & Anderson, 2012). The extensive training procedure employed here created strong memory traces, as evident by the overall high recognition rate on the final test. Therefore, even those memories that did not intrude are likely to have reached a moderate activation level in response to their cues, necessitating some engagement of control. However, when memories surpassed the activation threshold for recollection, the suppression mechanism might have had to be dynamically upregulated to purge the intruding memory. Whether or not awareness is necessary, our findings show that, when people attempt to exclude an intrusive memory from awareness, they appear to dynamically adjust a top–down mechanism that suppresses retrieval processes supported by the HC.

Efforts to suppress awareness of unwanted memories are ubiquitous in the aftermath of trauma, and involuntary memory intrusions constitute an important symptom present in many psychological disorders, including depression and post-traumatic stress disorder (Ehlers et al., 2002). To the extent that it disrupts traces, suppressing memories may reduce intrusions over time. Protracted engagement of suppression as a coping mechanism may, moreover, alter the potential for top–down modulation of HC in general. Indeed, in the year after a trauma, people undergo significant cortical plasticity, confined largely to DLPFC regions (Lyoo et al., 2011). Notably, individual variations in the increase in cortical thickness predicted the long-term recovery from post-traumatic stress disorder intrusion symptoms, and over time, cortical thickness returned to baseline with the degree that the patients experienced an alleviation of their symptoms. Although the mechanistic origins of this DLPFC plasticity and its function remain speculative, the present linkage between DLPFC engagement and the demand to terminate awareness provides an important candidate process. By identifying foundational neurocognitive mechanisms underlying people's coping response, the present findings thus introduce an important hypothesis about what alters people's mnemonic landscape and what permits adjustment after trauma. This discovery has translational potential for ameliorating the suffering of individuals for whom these mechanisms break down. More broadly, these findings point to a fundamental process by which we regulate the contents of our awareness.

Acknowledgments

We thank I. Charest and P. Gagnepain for assistance in data analysis and A. Gazzaley and J. Rissman for providing the pictures. This work was supported by the UK Medical Research Council (MC-A060-5PR00).

Reprint requests should be sent to Roland G. Benoit, Department of Psychology, Harvard University, Room 860, William James Hall, 33 Kirkland Street, Cambridge, MA 02138, or via e-mail: rbenoit@fas.harvard.edu.

REFERENCES

Aly, M., Knight, R. T., & Yonelinas, A. P. (2010). Faces are special but not too special: Spared face recognition in amnesia is based on familiarity. *Neuropsychologia, 48,* 3941–3948.

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature, 410,* 366–369.

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences, 18,* 279–292.

Anderson, M. C., & Huddleston, E. (2011). Towards a cognitive and neurobiological model of motivated forgetting. In R. F. Belli (Ed.), *True and false recovered memories: Toward a reconciliation of the debate*, Vol. 58, *Nebraska symposium on motivation* (pp. 53–120). New York: Springer.

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., et al. (2004). Neural systems underlying the suppression of unwanted memories. *Science, 303,* 232–235.

Benoit, R. G., & Anderson, M. C. (2012). Opposing mechanisms support the voluntary forgetting of unwanted memories. *Neuron, 670,* 450–460.

Benoit, R. G., Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2012). Rostral prefrontal cortex and the focus of attention in prospective memory. *Cerebral Cortex, 22,* 1876–1886.

Bergström, Z. M., de Fockert, J. W., & Richardson-Klavehn, A. (2009). ERP and behavioural evidence for direct suppression of unwanted memories. *Neuroimage, 48,* 726–737.

Bird, C. M., & Burgess, N. (2008). The hippocampus supports recognition memory for familiar words but not unfamiliar faces. *Current Biology, 18,* 1932–1936.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M., Dawson, D. R., et al. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society, 12,* 194–209.

Corallo, G., Sackur, J., Dehaene, S., & Sigman, M. (2008). Limits on introspection: Distorted subjective time during the dual-task bottleneck. *Psychological Science, 19,* 1110–1117.

Depue, B. E., Curran, T., & Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science, 317,* 215–219.

Detre, G. J., Natarajan, A., Gershman, S. J., & Norman, K. A. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia, 51,* 2371–2388.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psychology, 55,* 51–86.

Ehlers, A., Hackmann, A., Steil, R., Clohessy, S., Wenninger, K., & Winter, H. (2002). The nature of intrusive memories after trauma: The warning signal hypothesis. *Behaviour, Research & Therapy, 40,* 995–1002.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30,* 123–152.

Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience, 3,* 1149–1152.

Friston, K. J., Harrison, L., & Penny, W. D. (2003). Dynamic causal modelling. *Neuroimage, 19,* 1273–1302.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiack, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2,* 189–210.

Gagnepain, P., Henson, R., & Anderson, M. C. (2014). Suppressing unwanted memories reduces their unconscious influence via targeted cortical inhibition. *Proceedings of the National Academy of Sciences, U.S.A., 111,* 1310–1319.

Goldman-Rakic, P. S., Selemon, L. D., & Schwartz, M. L. (1984). Dual pathways connecting the dorsolateral prefrontal cortex with the hippocampal formation and parahippocampal cortex in the rhesus monkey. *Neuroscience, 12,* 719–743.

Hart, R. E., & Schooler, J. W. (2012). Suppression of novel stimuli: Changes in the accessibility of suppressed, nonverbalizable shapes. *Consciousness and Cognition, 21,* 1541–1546.

Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin & Review, 12,* 484–489.

Kim, D., Oh, D., Kim, S. H., Sim, K., & Lee, J. (2013). Effects of intentional suppression of recall of unwanted images in repressors and non-repressors. *Social Behavior and Personality, 41,* 319–326.

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience, 2,* 4.

Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience, 10,* 908–914.

Küpper, C. S., Benoit, R. G., Dalgleish, T., & Anderson, M. C. (2014). Direct suppression as a mechanism of controlling unpleasant memories in daily life. *Journal of Experimental Psychology: General, 143,* 1443–1449.

Levy, B. J., & Anderson, M. C. (2012). Purging of memories from conscious awareness tracked in the human brain. *The Journal of Neuroscience, 32,* 16785–16794.

Lyoo, I. K., Kim, J. E., Yoon, S. J., Hwang, J., Bae, S., & Kim, D. J. (2011). The neurobiological role of the dorsolateral prefrontal cortex in recovery from trauma. Longitudinal brain imaging study among survivors of the South Korean subway disaster. *Archives of General Psychiatry, 68,* 701–713.

Malian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage, 19,* 1233–1239.

Morris, R., Pandya, D. N., & Petrides, M. (1999). Fiber system linking the mid-dorsolateral frontal cortex with the retrosplenial/presubicular region in the rhesus monkey. *Journal of Comparative Neurology, 407,* 183–192.

Nader, K., Schafe, G. E., & Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature, 406,* 722–726.

Noreen, S., & MacLeod, M. D. (2013). It's all in the detail: Intentional forgetting of autobiographical memories using the autobiographical think/no-think task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 375–393.

Norman, K. A., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review, 114,* 887–953.

Paz-Alonso, P. M., Bunge, S. A., Anderson, M. C., & Ghetti, S. (2013). Strength of coupling within a mnemonic control network differentiates those who can and cannot suppress memory retrieval. *The Journal of Neuroscience, 33,* 5017–5026.

Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M., et al. (2010). Comparing families of dynamic causal models. *PLoS Computational Biology, 6,* e1000709.

Schacter, D. L., Guerin, S. A., & St Jacques, P. L. (2011). Memory distortion: An adaptive perspective. *Trends in Cognitive Sciences, 5,* 467–474.

Sergent, C., & Dehaene, S. (2004). Neural processes underlying conscious perception: Experimental findings and a global neuronal workspace framework. *The Journal of Physiology, 98,* 374–384.

Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E., Daunizeau, J., & Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *Neuroimage, 49,* 3099–3109.

Waldhauser, G., Lindgren, M., & Johansson, M. (2012). Intentional suppression can lead to a reduction of memory strength: Behavioral and electrophysiological findings. *Frontiers in Psychology, 3,* 401.

Wimber, M., Schott, B. H., Wendler, F., Seidenbecher, C. I., Behnisch, G., Macharadze, T., et al. (2011). Prefrontal dopamine and the dynamic control of human long-term memory. *Translational Psychiatry, 1,* e15.