Temporally discontinuous engagement of the vmPFC in remote memory retrieval

Daniel N. Barry[1], Martin J. Chadwick[2] and Eleanor A. Maguire[1*]

[1]Wellcome Centre for Human Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London, WC1N 3AR, UK

[2]Institute of Behavioural Neuroscience, Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, 26 Bedford Way, London, WC1H 0AP, UK

*Corresponding author: e.maguire@ucl.ac.uk

Competing interests: The authors declare that no competing interests exist.

Abstract

Systems-level consolidation is the time-dependent reorganisation of a memory trace in the neocortex, with the ventromedial prefrontal cortex (vmPFC) being particularly implicated. Capturing the precise temporal evolution of this crucial process in humans has long proved elusive. Here, we used multivariate methods and a longitudinal functional MRI design to detect, with high granularity, the extent to which autobiographical memories of different ages were represented in vmPFC and how this changed over time. We observed an unexpected biphasic involvement of vmPFC during retrieval, rising and falling around an initial peak of 8-12 months, before re-engaging for older two and five year old memories. Remarkably, when re-examined eight months later, representations of individual memories had undergone their hypothesised strengthening or weakening over time. We conclude that the temporal recruitment of vmPFC in autobiographical memory retrieval seems to be non-linear, revealing a previously-unknown feature of systems-level consolidation that is absent from current theories.

Introduction

Autobiographical memories are the cherished ghosts of our past. Through them we visit places long departed, see faces once familiar, hear voices now silent and re-experience emotions since dormant. These memories of our personal past experiences can be many decades old yet we are often able to recall them on a whim and with ease. How the brain represents autobiographical memories over a lifetime is one of the central, and as yet unanswered, questions of memory neuroscience.

Our fleeting present transitions into our autobiographical past through the modification of synaptic connectivity over the course of a few hours (Kandel, 2001), an undisputedly hippocampal-dependent process (Guzowski, 2002; Morris et al., 2003; Pastalkova et al., 2006; Runyan & Dash, 2005). A memory's journey does not end there, however, because over the course of time these memories come to be represented in the neocortex – this is termed systems-level consolidation (Frankland & Bontempi, 2005) – although the precise timeframe for this is unknown. Whether the hippocampus ever fully relinquishes its involvement in autobiographical memory retrieval is a long-standing debate. One theory asserts that the hippocampus plays a short-term role before memories become fully consolidated to the neocortex and can be retrieved without the hippocampus (Squire, Genzel, Wixted, & Morris, 2015). Other accounts posit that the hippocampus is necessary for the vivid retrieval of autobiographical memories in perpetuity (Maguire, 2014; Maguire & Mullally, 2013; Moscovitch, Cabeza, Winocur, & Nadel, 2016; Moscovitch et al., 2005; Winocur & Moscovitch, 2011).

By contrast, there is universal agreement that autobiographical memories are represented in the neocortex at some point in time. Consequently, therefore, scrutinising the neocortical targets of systems-level consolidation may offer a clearer view of the path an autobiographical memory takes over time. While the neocortical areas into which autobiographical memories are consolidated are not always specified in theoretical accounts, animal experiments have consistently implicated the

medial prefrontal cortex. Although some evidence has linked this region with the formation (Lesburgueres et al., 2011; Zhao et al., 2005) and recall of recently acquired memories (Einarsson & Nader, 2012; Leon, Bruno, Allard, Nader, & Cuello, 2010; Tse et al., 2011), it seems to be disproportionately involved in the retrieval of memories learned weeks previously (Ding, Teixeira, & Frankland, 2008; Frankland, Bontempi, Talton, Kaczmarek, & Silva, 2004; Kitamura et al., 2017; Liu, Zheng, & Li, 2009; Lopez et al., 2012; Luo et al., 2017; Maviel, Durkin, Menzaghi, & Bontempi, 2004; Takehara, Kawahara, & Kirino, 2003; Teixeira, Pomedli, Maei, Kee, & Frankland, 2006). Furthermore, recall success depends on post-learning activation (Takehara-Nishiuchi, Nakao, Kawahara, Matsuki, & Kirino, 2006) and structural changes in this region over time (Bero et al., 2014; Restivo, Vetere, Bontempi, & Ammassari-Teule, 2009; Vetere et al., 2011).

In humans, a putative functional homologue in the context of autobiographical memory is the ventromedial prefrontal cortex (vmPFC). Damage to this region in humans has been linked to impoverished recall of recent and remote autobiographical memories (e.g., Bertossi, Tesini, Cappelli, and Ciaramelli, 2016). In a recent review, however, McCormick, Ciaramelli, De Luca, and Maguire (2017) noted that it is difficult to come to a firm conclusion about the status of autobiographical memory in vmPFC-damaged patients. This is due to the dearth of studies examining autobiographical memory retrieval in detail in patients with selective bilateral vmPFC damage. Large, non-selective lesions, differences in how memories are elicited and cued, and confabulation (the oblivious recollection of blatant untruths; Turner, Cipolotti, Yousry, & Shallice, 2008) that often accompanies vmPFC pathology, make testing autobiographical memory retrieval a neuropsychological challenge.

Delay-dependent increases in vmPFC activity during memory recall have been demonstrated in some standard functional MRI (fMRI) studies of healthy individuals (Takashima et al., 2009; Takashima et al., 2006), but not others (Furman, Mendelsohn, & Dudai, 2012; Harand et al., 2012; Watanabe et al.,

2012). Likewise, while autobiographical memory retrieval induces reliable and robust engagement of the vmPFC (Svoboda, McKinnon, & Levine, 2006), it is unclear whether this activity increases (Niki & Luo, 2002), decreases (Oddo et al., 2010), or remains constant in accordance with memory remoteness (Gilboa, Winocur, Grady, Hevenor, & Moscovitch, 2004; Maguire & Frith, 2003; Maguire, Henson, Mummery, & Frith, 2001; Piefke, Weiss, Markowitsch, & Fink, 2005; Piolino et al., 2004; Rekkas & Constable, 2005; Ryan et al., 2001; Soderlund, Moscovitch, Kumar, Mandic, & Levine, 2012; Steinvorth, Corkin, & Halgren, 2006; Tsukiura et al., 2002).

Multi-voxel pattern analysis (MVPA) methods as applied to fMRI data are better positioned to bridge the empirical gap between the human and animal literatures due to their increased sensitivity to representations of specific neural patterns (Chadwick, Bonnici, & Maguire, 2012). Using an MVPA approach, Bonnici et al. (2012) demonstrated that remote 10 year old autobiographical memories were more detectable in the vmPFC than recent two week old autobiographical memories, consistent with its proposed role as a long-term consolidation site. This difference was not apparent in other cortical areas. In a follow up study two years later with the same participants and memories, the previously two week old autobiographical memories were now two years old and equally detectable in the vmPFC as the remote memories (Bonnici & Maguire, 2017). This suggests that systems-level consolidation of an autobiographical memory in the vmPFC is accomplished by at most two years, and perhaps even sooner.

If we are to gain traction on the important question of the time course of autobiographical memory consolidation in the vmPFC then clearly this wide two year time envelope needs to be more precisely resolved. This is what we aimed to do in the current fMRI study where we first compared the neural representations of autobiographical memories in the vmPFC sampled at four month intervals spanning a two year period. We then sought to verify any apparent time related differences in the representation of these memories by capturing their further neural evolution in a follow up study

involving the same participants and memories eight months later. This challenging longitudinal design in combination with multivariate methods provided an unprecedented level of granularity in characterising the dynamics of autobiographical memory consolidation in the vmPFC. We considered at least two alternative outcomes to be plausible. Either the representation of autobiographical memories in vmPFC might be gradual and linear depending on a memory's age, or there might be a step change in the detectability of an autobiographical memory in vmPFC perhaps in the early months post-encoding.

Results

Experiment 1

One week prior to the fMRI scan, with the assistance of personal photographs, participants (n=30) verbally recalled and rated the characteristics of autobiographical memories from eight time periods: memories that were 0.5 months old (0.5M, i.e., two week old memories), 4M, 8M, 12M, 16M, 20M, 24M and also 60M old – these latter memories being included as a benchmark for definitely remote (5 year old) memories (see Materials and Methods, Figure 1A). Two memories from each time period which were sufficiently vivid, detailed, specific and unique in time and place were chosen for subsequent recall in the scanner. This meant that there were two full sets of memories. Participants created a short phrase pertaining to an autobiographical memory, which was paired with the photograph to facilitate recall during the subsequent fMRI scan.

Subjective ratings of memory characteristics

While all memories satisfied the criteria of being vivid and detailed, and the ratings were high (Figure 1; see means and SDs in Supplemental Table 1A), vividness nevertheless varied as a function of memory age ($F_{(7,203)} = 3.45$, $p = 0.002$), with the most recent, 0.5M old, memories rated higher than 12M ($t_{29} = 4.08$, $p = 0.009$), 20M ($t_{29} = 3.88$, $p = 0.016$), 24M ($t_{29} = 4.18$, $p = 0.007$) and 60M old memories ($t_{29} = 3.45$, $p = 0.049$, Figure 1B). Subjective ratings of detail also differed across time-

points ($F_{(7,203)}$ = 5.74, p < 0.001), once again the most recent 0.5M old memories were rated higher than 4M (t_{29} = 4.45, p = 0.003), 8M (t_{29} = 3.97, p = 0.012), 12M (t_{29} = 5.00, p < 0.001), 16M (t_{29} = 4.96, p < 0.001), 20M (t_{29} = 5.37, p < 0.001), 24M (t_{29} = 4.51, p = 0.003) and 60M old memories (t_{29} = 3.98, p = 0.012, Figure 1C). The expenditure of effort during recall also varied according to remoteness of memories ($F_{(7,203)}$ = 5.79, p < 0.001), with 0.5M old memories being easier to recollect than 12M (t_{29} = -5.29, p < 0.001), 16M (t_{29} = -3.90, p = 0.015), 20M (t_{29} = -3.67, p = 0.027) and 60M old memories (t_{29} = -4.55, p = 0.003, Figure 1D). No significant difference was observed across time periods from 4M to 60M on any of these characteristics (all p > 0.05), nor did memories differ in their personal significance ($F_{(7,203)}$ = 1.66, p = 0.120, Figure 1E) or emotional valence ($F_{(7,203)}$ = 1.51, p = 0.166, Figure 1F) as a function of age.

In addition to these main ratings of interest, no difference was reported in the extent to which memories were recalled as active or static ($F_{(7,203)}$ = 1.36, p = 0.224), or from a first or third person perspective ($F_{(3.69,107.02)}$ = 1.09, p = 0.365) across time periods. The reported frequency at which memories were recalled since the original event (rated on a five point scale from “never” to “very frequently”), differed as a function of time ($F_{(5.11,148.04)}$ = 4.36, p < 0.001), with the most recent 0.5M old memories thought about more frequently than 12M (t_{29} = 4.37 p = 0.004), 16M (t_{29} = 3.47, p = 0.046) and 24M (t_{29} = 3.71, p = 0.024) old memories.

Overall, therefore, memories were generally well matched on subjective phenomenological ratings, satisfied the criteria of high quality of memory recall, with only small differences observed for the most recent 0.5M old memories compared to the other autobiographical memories, as might be expected.

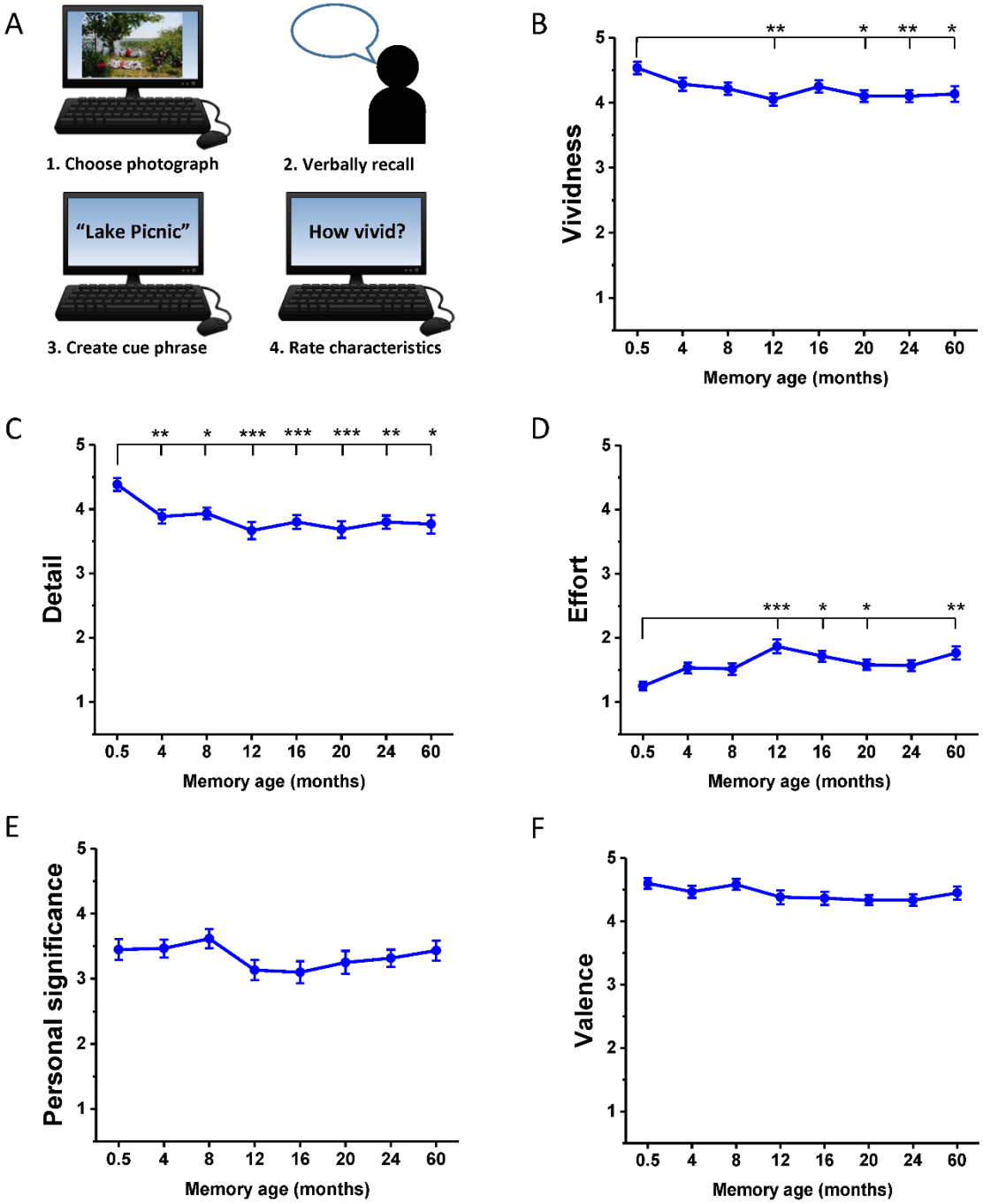

Figure 1. Memory harvesting and subjective ratings. (**A**) Schematic of the interview where the autobiographical memories were harvested. Participants recalled a memory cued by a personal photograph, chose a phrase to help remind them of this memory during the subsequent scanner task, and rated its characteristics. (**B-F**) Subjective ratings (means +/- 1SEM; see also Supplemental Table 1A) of memory characteristics at each time period for Experiment 1, averaged across the two sets of memories. Ratings were on a scale of 1 to 5, where 1 was low and 5 was high. For emotional valence: 1-2 = negative, 3 = neutral, 4-5 = positive. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Objective scoring of memory details

To complement the subjective ratings of memory characteristics with a more objective assessment of their content, transcripts of participants' memory interviews were scored using the Autobiographical Interview protocol (Levine, Svoboda, Hay, Winocur, & Moscovitch, 2002; Materials and Methods). In total for this first experiment, 10,187 details were scored. The mean (SD) number of internal details (bound to the specific 'episodic' spatiotemporal context of the event) and external details (arising from a general 'semantic' knowledge or references to unrelated events) are shown in Supplemental Table 1B (see also Figure 2). They were then compared across time periods. In contrast to the subjective ratings of memory detail, there was no difference in the number of details recalled across memories from different time periods ($F_{(4.54,131.66)} = 1.92$, $p = 0.101$). As expected, the number of internal and external details differed ($F_{(1,29)} = 206.03$, $p < 0.001$), with more internal details recalled for every time period (all $p < 0.001$). No interaction between time period and type of detail was observed ($F_{(7,203)} = 1.87$, $p = 0.077$). The number of external details recalled was remarkably consistent across all time periods, emphasising the episodic nature of recalled events irrespective of remoteness. Inter-rater reliabilities for the scoring (see Materials and Methods) were high for both internal (ICC = 0.94) and external (ICC = 0.81) details.

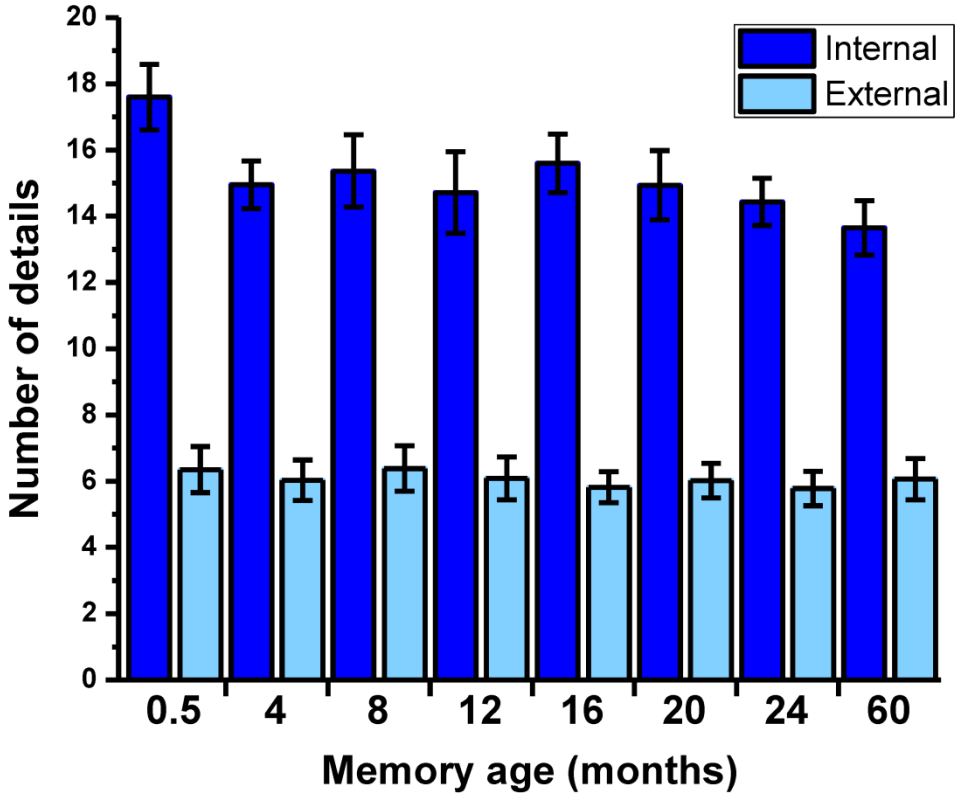

Figure 2. Objective analysis of memory details. The mean +/- 1SEM (see also Supplemental Table 1B) number of internal and external details at each time period, averaged across the two sets of autobiographical memories.

Memory representations in the vmPFC

Ventromedial prefrontal cortex was delineated as the ventral medial surface of the frontal lobe and the medial portion of the orbital frontal cortex (Mackey & Petrides, 2014). This comprises areas implicated in memory consolidation (Bonnici et al., 2012; Bonnici & Maguire, 2017; Nieuwenhuis & Takashima, 2011), namely Brodmann Areas 14, 25, ventral parts of 24 and 32, the caudal part of 10 and the medial part of BA 11 (see Figure 3A, and Materials and Methods).

On each trial, the photograph and associated pre-selected cue phrase relating to each event were displayed on a screen for 3 seconds. Following removal of this cue, participants then closed their eyes and recalled the memory. After 12 seconds, the black screen flashed white twice, to cue the participant to open their eyes. The participant was then asked to rate how vivid the memory recall

had been using a five-key button box, on a scale of 1-5, where 1 was not vivid at all, and 5 was highly vivid (see Figure 3B).

We used Representational Similarity Analysis (RSA) to quantify the extent to which the strength of memory representations in the vmPFC differed as a function of memory age. This was achieved by contrasting the similarity of neural patterns when recalling the same memory with their similarity to other memories to yield a "neural representation" score for each memory (see Materials and Methods and Figure 3C). As there were two memories recalled per time period, the neural representation scores were averaged to produce one value for that period.

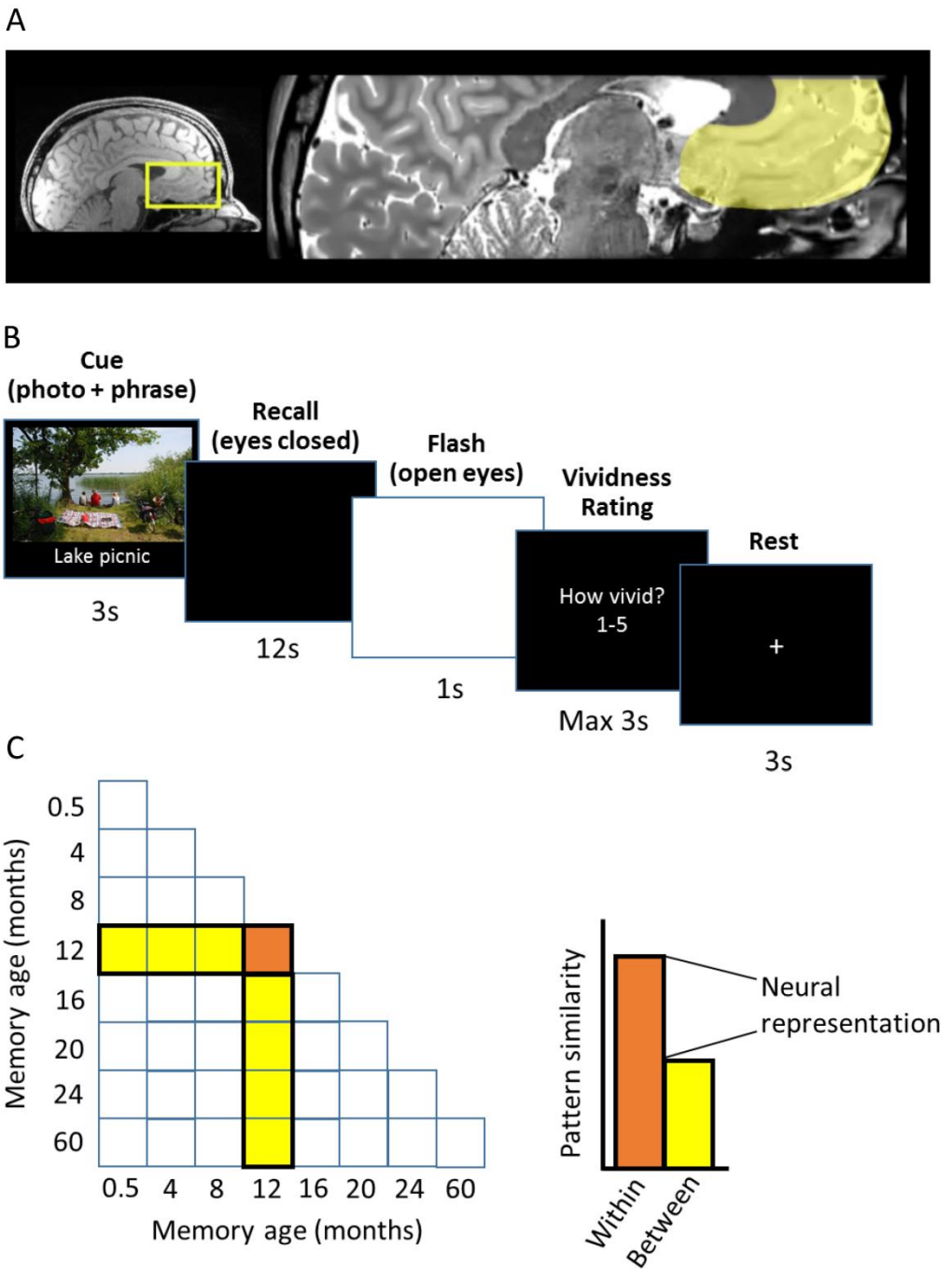

Figure 3. Experimental details. (**A**) The vmPFC highlighted on an example participant's structural MRI scan. (**B**) Timeline of an example trial from the scanning task. (**C**) Graphical illustration of the neural representation score calculation using RSA. The neural pattern similarity across trials recalling the same memory (orange) minus the mean pattern similarity between that memory and other memories (yellow) generates a "neural representation" score. A score significantly higher than zero indicates a neural pattern distinct to that memory is present in the vmPFC.

We anticipated an increase in the strength of memory representations at some point between 0.5M and 24M, in line with the results of Bonnici et al. (2012) and Bonnici and Maguire (2017). This is what we observed, where the most recent 0.5M memories were undetectable (t_{29} = 0.72, p = 0.477) in vmPFC, in contrast to the distinct neural signatures observed for 4M (t_{29} = 2.85, p = 0.008), 8M (t_{29} = 3.09, p = 0.004) and 12M (t_{29} = 3.66, p < 0.001) old memories (see Figure 4A). These changes in the strength of memory representations were significant across time periods ($F_{(7,203)}$ = 2.22, p = 0.034), with an observed increase in vmPFC recruitment from 0.5M to 8M (t_{29} = 2.07, p = 0.048) and 12M (t_{29} = -2.20, p = 0.036). However, what was observed for the following two time periods was unexpected – an apparent disengagement of the vmPFC over the next eight months as we observed weak detectability of memory representations in vmPFC for 16M (t_{29} = 1.85, p = 0.074) and 20M (t_{29} = 1.03, p = 0.310) old memories. Neither 16M (t_{29} = -1.06, p = 0.298) nor 20M memories (t_{29} = -0.40, p = 0.691) were more strongly represented than the recent 0.5M old memories. In contrast, the more remote 24M (t_{29} = 4.34, p < 0.001) and 60M (t_{29} = 3.55, p = 0.001) memories were detectable in the vmPFC, and significantly more so than the most recent memories (24M vs 0.5M, t_{29} = -2.93, p = 0.007; 60M vs 0.5M, t_{29} = -2.54, p = 0.017) as well as the more temporally proximal 20M old memories (24M vs 20M, t_{29} = -2.50, p = 0.018; 60M vs 20M, t_{29} = -2.32, p = 0.028).

The experimental design afforded us the opportunity to verify this biphasic pattern. As we sampled two memories per time-point, this time-dependent pattern should be evident in both sets of memories. As shown in Figure 4B, the two sets of memories followed a similar time-course of changes in representation within vmPFC. This is a compelling replication, given that the two memories from each time-period were unrelated in content as a prerequisite for selection, recalled in separate sessions in the scanner and analysed independently from each other.

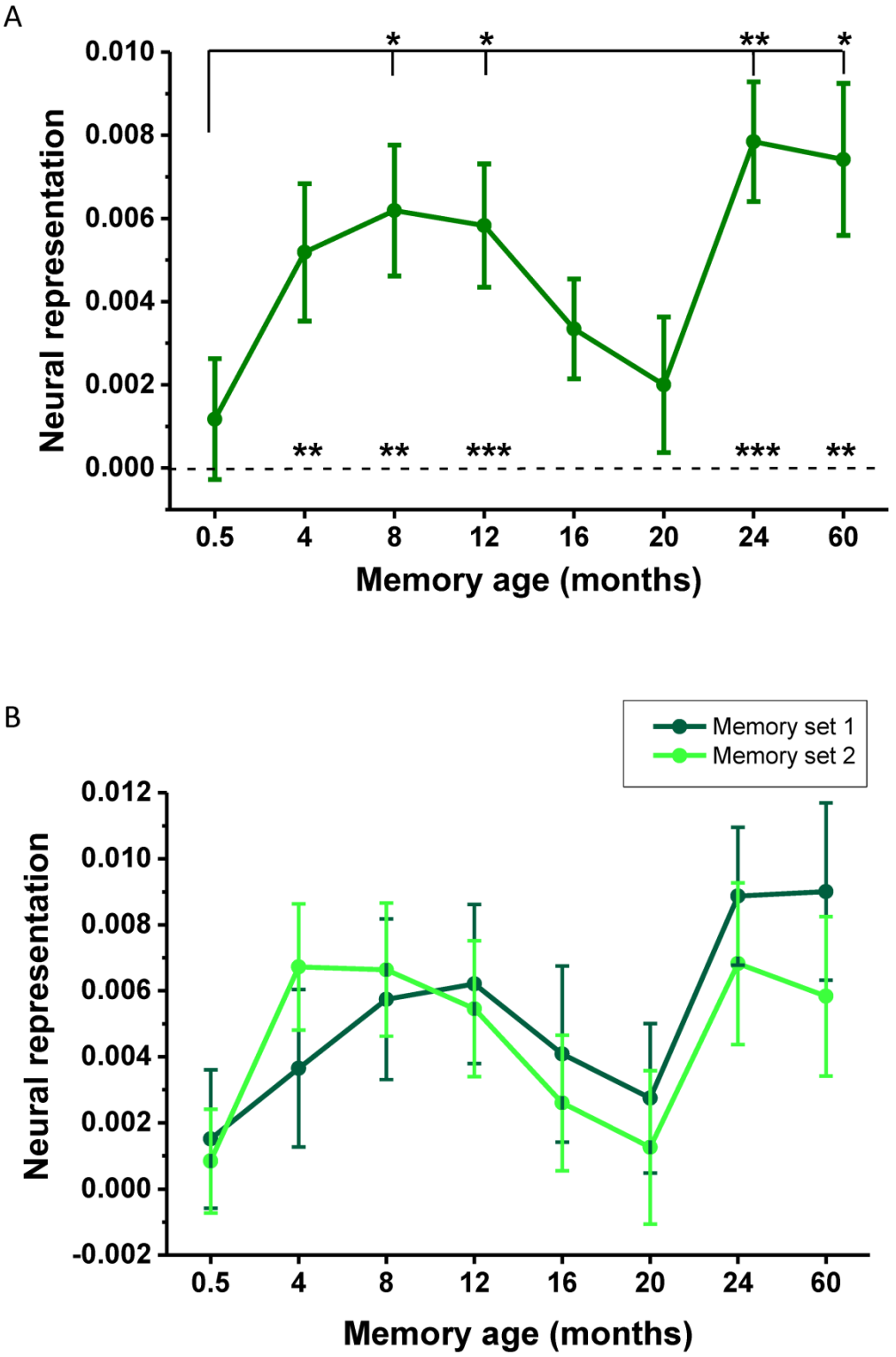

Figure 4. fMRI results of Experiment 1. (**A**) Mean +/- 1SEM neural representation scores at each time-point averaged across the two sets of memories. Asterisks above the dotted line indicate detectability of memories in vmPFC at each time-point. Asterisks above the solid line indicate significant increases in memory representations from the most recent (0.5M old) memories. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (**B**) Neural representation scores at each time-point plotted separately for the two sets of autobiographical memories.

Our main focus was the vmPFC, given previous work highlighting specifically this region's role in representing autobiographical memories over time (Bonnici et al., 2012; Bonnici & Maguire, 2017). We also scanned within a partial volume, so were constrained in what other brain areas were available for testing (see Materials and Methods). Nevertheless, we examined the same areas as Bonnici et al. (2012) and Bonnici and Maguire (2017), and in no case did we observe a significant change in memory detectability across time periods in the entorhinal/perirhinal cortex ($F_{(7,203)} = 1.55$, $p = 0.154$), hippocampus ($F_{(7,203)} = 0.98$, $p = 0.445$), posterior parahippocampal cortex ($F_{(7,203)} = 1.41$ $p = 0.202$), retrosplenial cortex ($F_{(7,203)} = 0.74$, $p = 0.641$), temporal pole ($F_{(7,203)} = 1.78$, $p = 0.093$), lateral temporal cortex ($F_{(4.86,141.03)} = 0.68$, $p = 0.636$) or lateral visual cortex ($F_{(7,203)} = 0.96$, $p = 0.465$).

Following scanning, participants completed three additional ratings. They were asked to indicate the extent to which the memories were changed by the 6 repetitions during scanning on a scale ranging from 1 (not at all) to 5 (completely). They reported that the memories were not changed very much by repetition (mean: 2.61, SD: 0.74). They were also asked how often during scanning they thought about the memory interview one week previous on a scale of 1 (not at all) to 5 (completely), with participants indicating they rarely thought about the interview (mean: 2.29, SD: 1.01). Finally, participants were asked the extent to which the recall of memories from each time period unfolded in a consistent manner over the course of the session. A difference was observed ($F_{(7,203)} = 2.78$, $p = 0.009$), with the most recent 0.5M old memories being rated as more consistently recalled than the most remote 60M memories ($t_{29} = 3.97$, $p = 0.012$).

Rationale and predictions for Experiment 2

The biphasic pattern we observed in the fMRI data did not manifest itself in the subjective or objective behavioural data. In fact, the only difference in those data was higher ratings for the most recent 0.5M old memories. However, these were paradoxically the most weakly represented

memories in the vmPFC, meaning the neural patterns were not driven by memory quality. The objective scoring of the memories confirmed comparable levels of detail provided for all memories, without any significant drop in episodic detail or increase in the amount of semantic information provided as a function of time. Therefore, the amount or nature of the memory details were not contributing factors.

Nevertheless, to verify that the results genuinely represented the neural correlates of memory purely as a function of age, one would need to study the effects of the passage of time on the individual neural representations. Therefore we invited the participants to revisit eight months later to recall the same memories again both overtly and during scanning; 16 of the participants agreed to return. In order to generate specific predictions for the neural representations during Experiment 2, we took the actual data for the 16 subjects from Experiment 1 who returned eight months later (see Figure 5 green line, where the biphasic pattern is still clearly evident), and shifted it forwards by two time-points to simulate the expected pattern eight months later (Figure 5 grey line). Note that for the 28M and 32M time periods in Experiment 2 we assumed they would have the same level of detectability as 24M old memories given the absence of data relating to these time periods from Experiment 1. We further assumed the neural representations between 60M and 68M would be unchanged.

A significant difference between original and predicted neural representation scores from any time period would generate a hypothesised change. Accordingly, 0.5M old memories were hypothesised to be more detectable eight months later (t_{15} = -2.85, p = 0.012), while the original 4M (t_{15} = -0.40, p = 0.695) and 8M (t_{15} = 0.80, p = 0.436) old memories should remain unchanged. Twelve month old memories from Experiment 1 should decrease in detectability (t_{15} = 2.61, p = 0.020), whereas 16M old memories should not differ significantly in their representations at 24M (t_{15} = -1.53, p = 0.146). Original 20M old memories should be better represented at 28M (t_{15} = -4.15, p < 0.001). Finally, the

original 24 and 60 month memories were not assumed to change over time in the strength of neural representations. Overall, therefore, while an increase in detectability in vmPFC of the 0.5M memories eight months later is an obvious prediction, the unexpected predictions generated by the Experiment 1 data were a decrease in detectability of the previously well-represented 12M old memories and an increase in the detectability of the previously undetectable 20M old memories, with no concomitant changes in the behavioural data.

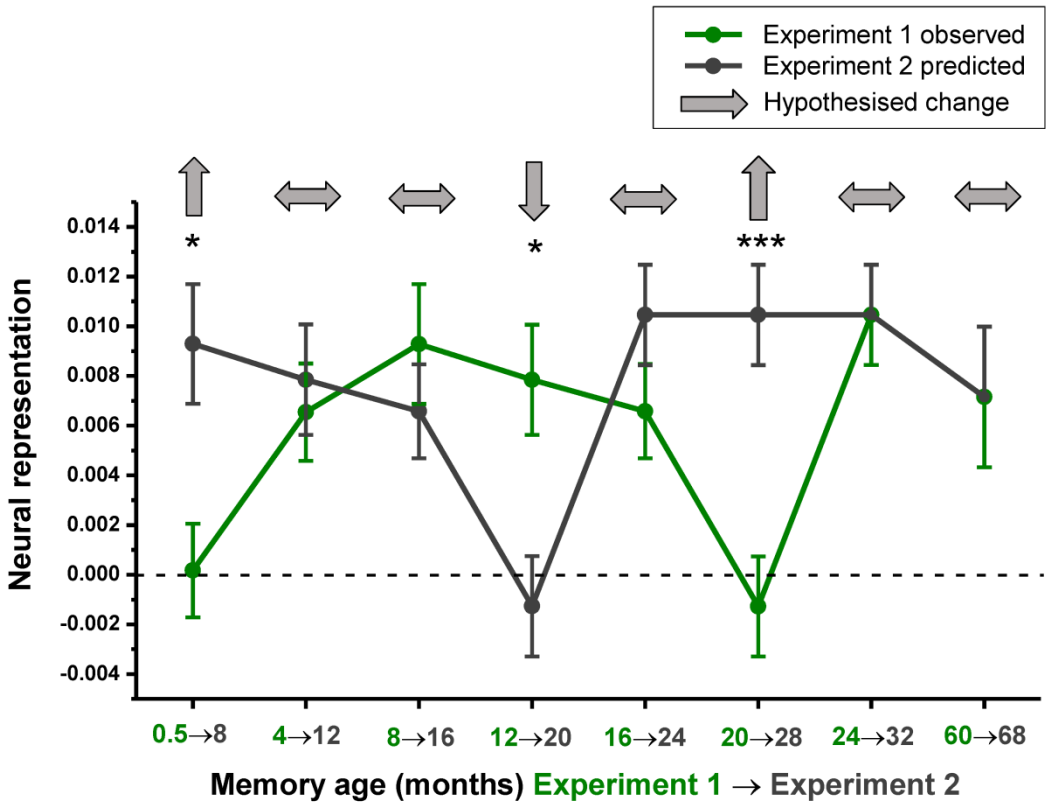

Figure 5. Predicted fMRI changes eight months later in Experiment 2. Predicted mean +/- 1SEM changes in the neural representations of individual autobiographical memories after eight months (dark grey line), based on shifting the original observed data forward by two time-points for the 16 subjects from Experiment 1 (green line) who returned for Experiment 2. Light grey arrows indicate the hypotheses. * $p < 0.05$, ** $p < 0.01$.

Experiment 2 (eight months later)

One week prior to the fMRI scan, with the assistance of the personal photographs and previously chosen phrases which were used as cues in Experiment 1, the participants verbally recalled and

rated the characteristics of their autobiographical memories just as they had done eight months previously (see Materials and Methods and Figure 6A).

Subjective ratings of memory characteristics

Means and SDs are provided in Supplemental Table 2A. Autobiographical memories recalled during Experiment 2 did not differ across time periods on vividness ($F_{(7,105)} = 0.83$, $p = 0.564$), detail ($F_{(7,105)} = 1.30$, $p = 0.257$), effort ($F_{(7,105)} = 0.11$, $p = 0.998$), personal significance ($F_{(7,105)} = 1.49$, $p = 0.180$), valence ($F_{(7,105)} = 1.06$, $p = 0.397$), viewpoint ($F_{(3.42,51.22)} = 1.24$, $p = 0.31$) or motion ($F_{(3.95,59.32)} = 1.43$, $p = 0.237$). When asked how frequently they had thought about the autobiographical memories in the eight months between experiments (rated on a five point scale from "never" to "very frequently"), participants reported some change across time periods ($F_{(7,105)} = 3.04$, $p = 0.006$). However, the only significant difference between time periods was a lower recall frequency for now 32M old memories compared to the now 12M ($t_{15} = 3.87$, $p = 0.042$). Given the range of responses to this question across conditions (1.50-2.03), clearly participants had not given the memories much thought in the intervening eight months. Therefore, all memories recalled in Experiment 2 were extremely well matched in terms of their phenomenology, which reflects the consistency observed in ratings from eight months onwards in Experiment 1.

There were, however, differences in the absolute values of subjective ratings between experiments. There was a decrease in the reported vividness of all memories from Experiment 1 to Experiment 2 ($F_{(1,15)} = 88.45$, $p < 0.001$), from 0.5M to when they were 8M old ($t_{15} = 6.21$, $p < 0.001$), 8M to 16M ($t_{15} = 4.21$, $p = 0.006$), 12M to 20M ($t_{15} = 5.48$, $p < 0.001$), 16M to 24M ($t_{15} = 7.07$, $p < 0.001$), 20M to 28M ($t_{15} = 4.10$, $p = 0.008$), 24M to 32M ($t_{15} = 5.97$, $p < 0.001$) and 60M to 68M ($t_{15} = 5.33$, $p < 0.001$; Figure 6B). A comparable change was observed in the subjective impression of memory detail recalled following the eight month interlude ($F_{(1,15)} = 126.81$, $p < 0.001$), with a drop from 0.5M to 8M ($t_{15} = 6.26$, $p < 0.001$), 8M to 16M ($t_{15} = 4.03$, $p = 0.009$), 12M to 20M ($t_{15} = 4.78$, $p = 0.002$), 16M

to 24M (t_{15} = 3.72, p = 0.016), 20M to 28M (t_{15} = 3.67, p = 0.018), 24M to 32M (t_{15} = 4.55, p < 0.003) and 60M to 68M (t_{15} = 9.67, p < 0.001; Figure 6C). Recalling memories eight months later was also perceived as more effortful ($F_{(1,15)}$ = 43.32, p < 0.001), from 0.5M to 8M (t_{15} = -7.81, p < 0.001), 4M to 12M (t_{15} = -3.30, p = 0.039), 16M to 24M (t_{15} = -1.95, p = 0.021), and 20M to 28M (t_{15} = -4.03, p = 0.009; Figure 6D). The elapsed time between experiments also led to a reduction in the reported personal significance of memories ($F_{(1,15)}$ = 11.82, p = 0.004), from 24M to 32M (t_{15} = 3.58, p = 0.022; Figure 6E). Ratings of emotional valence also changed over the eight month period ($F_{(1,15)}$ = 9.78, p = 0.007), with a reported attenuation of the positivity of memories from 12M to 20M (t_{15} = 3.87, p = 0.012; Figure 6F). In addition to these main ratings, no difference was reported in the extent to which memories were recalled from a first or third person perspective ($F_{(1,15)}$ = 0.513, p = 0.485) over the eight month period. The extent to which memories were recalled as active or static was altered by the passage of time between experiments ($F_{(1,15)}$ = 11.01, p = 0.005), with the original 0.5M old memories becoming more static when 8M old (t_{15} = -3.42, p = 0.031).

Despite the observed changes in some subjective ratings from Experiment 1 to Experiment 2, they were unidirectional across all time periods. As such, if the pattern of hypothesised emergence and disappearance of neural representations in vmPFC were to be supported in Experiment 2, then it could not be accounted for by changes in subjective ratings. Additionally, although the changes in subjective ratings across time tend to suggest a comparable degradation in memory quality across all time periods, this may be misleading. These absolute changes in values could be influenced by participants' expectations of their ability to recall memories after an extended period of time with high fidelity, because the objective scoring of memory detail revealed no such pattern, as we report in the next section.

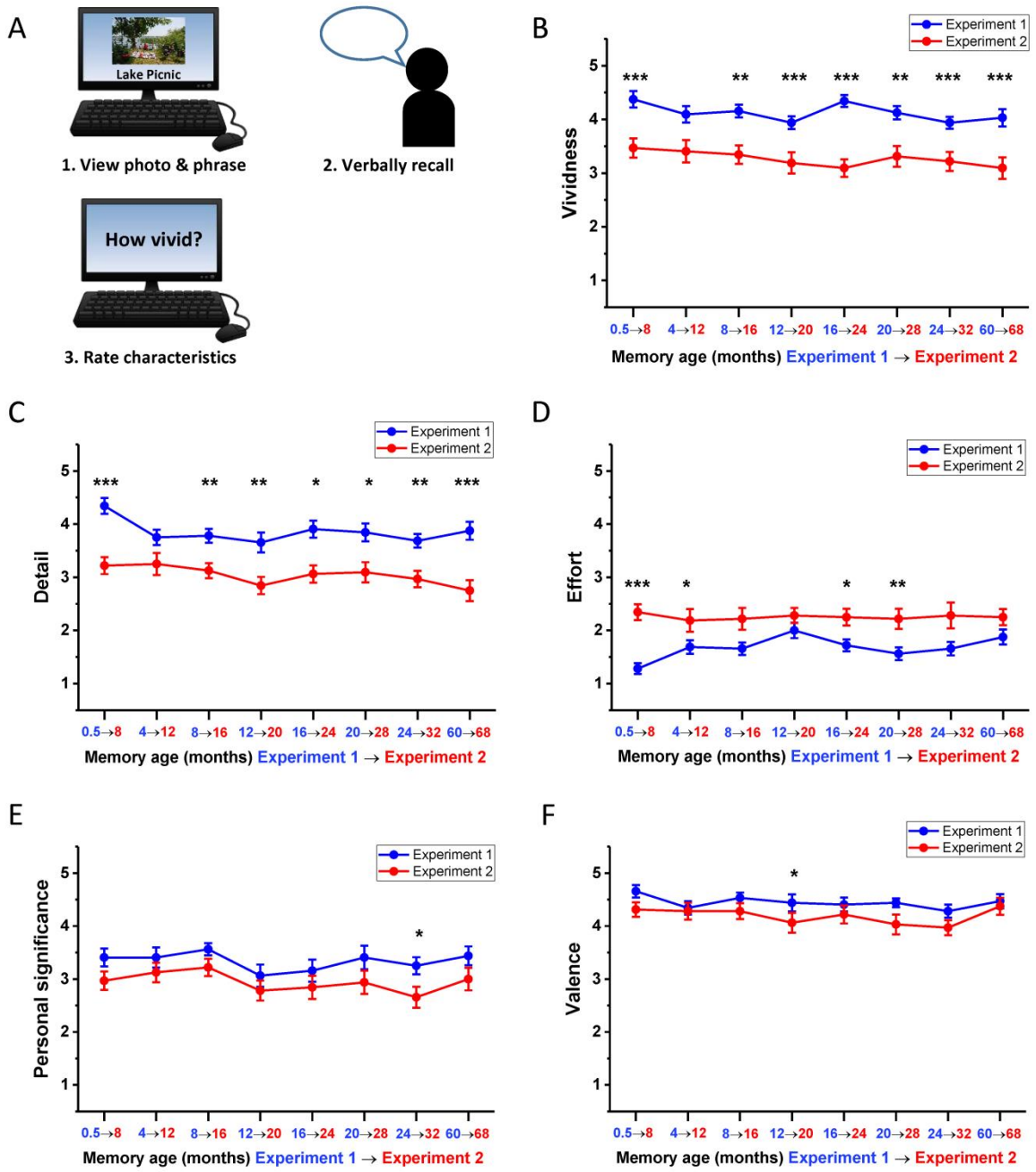

Figure 6. Memory recall and subjective ratings. (**A**) Schematic of the interview where participants recalled an autobiographical memory using their previously chosen photograph and cue phrase and rated its characteristics. (**B-F**) Subjective ratings (means +/- 1SEM; see also Supplemental Table 2A) of memory characteristics at each time period for Experiment 1 (blue line, n=16) and how the ratings of the same memories differed eight months later during Experiment 2 (red line, n=16) averaged across the two sets of memories in both cases. Ratings were on a scale of 1 to 5, where 1 was low and 5 was high. For emotional valence: 1-2 = negative, 3 = neutral, 4-5 = positive. Asterisks indicate significant differences in memory ratings between Experiment 1 and 2; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Objective scoring of memory details

As with Experiment 1, transcripts of participants' memory interviews during Experiment 2 were scored using the Autobiographical Interview protocol (Levine et al., 2002; see Materials and Methods). A total of 6,444 details were scored (see Supplemental Table 2B for means, SD). There was a difference in the number of details recalled across different time periods in Experiment 2 ($F_{(7,105)}$ = 2.49, p = 0.021). However, this difference was only observed for external details ($F_{(7,105)}$ = 3.25, p = 0.004), with more provided for 28M memories than 12M memories (t_{15} = -4.68, p = 0.008). As with Experiment 1, the number of internal and external details differed ($F_{(1,15)}$ = 72.57, p < 0.001), with more internal details recalled for every time period (all p < 0.01). No interaction between time period and type of detail was observed ($F_{(7,105)}$ = 0.87, p = 0.530).

When the objective scores for both experiments were compared, no significant difference was observed in the overall number of details provided eight months later ($F_{(1,15)}$ = 1.93, p = 0.185; see Figure 7). Furthermore, there was no significant interaction effect between experiment and time period ($F_{(1,15)}$ = 1.97, p = 0.066), indicating that the amount of details provided for memories from any particular time period in Experiment 1 were not affected by the passage of time. Finally, no interaction effect was observed between experiment and type of detail provided ($F_{(1,15)}$ = 2.27, p = 0.153), showing that the ratio of internal to external details was preserved across experiments.

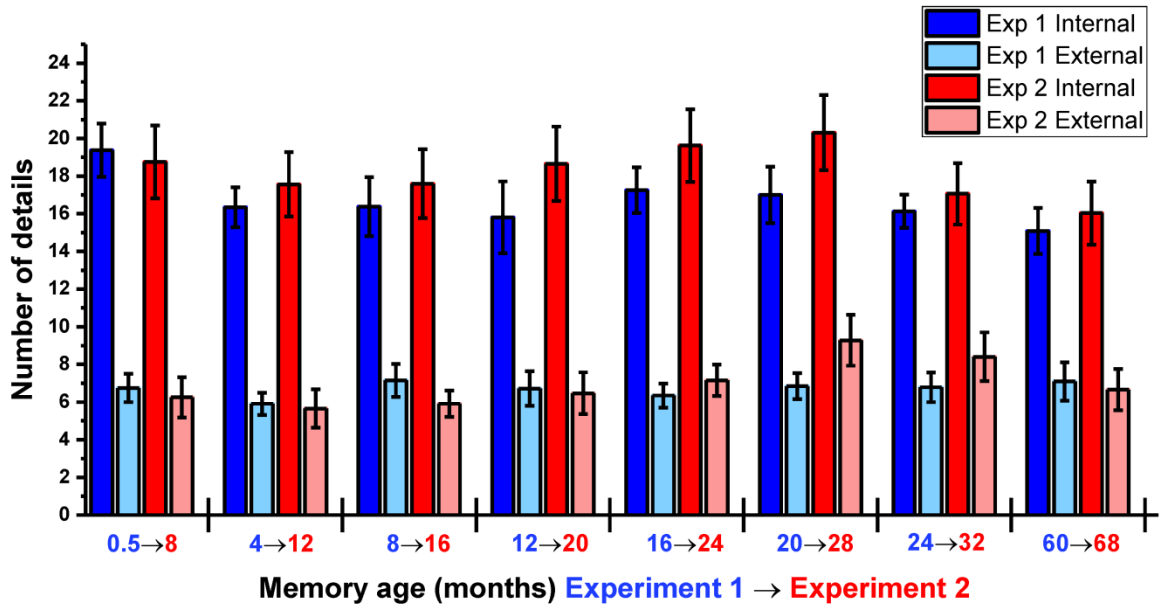

Figure 7. Objective analysis of memory details over time. The mean +/- 1SEM (see also Supplemental Table 2B) number of internal and external details at each time period for Experiment 1 (blue bars, n=16) and Experiment 2 (red bars, n=16), averaged across the two sets of autobiographical memories.

Changes in memory representations in the vmPFC

Participants were scanned in an identical fashion as Experiment 1 (see Materials and Methods and Figure 3B), and neural representation scores for memories from each time point were again calculated.

When comparing the neural representation scores of memories from the eight original time periods in Experiment 1 with those of the same memories eight months later during Experiment 2, a main effect for experiment ($F_{(1,15)}$ = 2.35, p = 0.146), or time period ($F_{(7,105)}$ = 1.18, p = 0.323), was not observed, however, an interaction between experiment and time period emerged ($F_{(7,105)}$ = 3.46, p = 0.002). Closer examination via our planned comparisons (see Figure 5 for a reminder of our predictions) revealed that seven out of the eight predictions made on the basis of the Experiment 1 findings were supported (Figure 8A). The original 0.5M old memories had increased in their representational strength in vmPFC eight months later (t_{15} = -1.84, p = 0.043), while the neural

representation scores of the 4M and 8M old memories were essentially unchanged at 12M (t_{15} = 0.43, p = 0.677) and 16M (t_{15} = 1.22, p = 0.242) respectively. As expected, the original 12M old memories from Experiment 1 were eight months later more poorly represented in vmPFC when 20M months old (t_{15} = 1.85, p = 0.042). The original 16M old memories were unchanged in their representational strength at 24M (t_{15} = 1.38, p = 0.187), while 20M old memories were significantly more detectable in vmPFC at 28M (t_{15} = -2.69, p = 0.008). The most remote 60M memories did not differ in their neural representation scores eight months later (t_{15} = 0.86, p = 0.402). In fact the only finding which was inconsistent with the predictions generated by Experiment 1 was a decrease in the representation of 24M old memories when they were 32M of age (t_{15} = -2.69, p = 0.009). However, this prediction was based on the assumption that memories do not undergo further dynamic shifts in neural representation between two and five years, which may not be the case, and we did not have 32M data from Experiment 1 to corroborate this finding.

For completeness, Figure 8B plots the neural representation scores for the two sets of memories in Experiment 2. As previously observed in Experiment 1, the two sets of memories displayed a similar time-course in terms of their neural representations, despite being recalled in separate scanning sessions, in a randomised order and analysed separately.

As with Experiment 1, when examining other areas within the partial volume, in no case did we find a significant difference in memory detectability across time periods.

Following scanning, participants completed three additional ratings. They were asked to indicate the extent to which the memories were changed by the 6 repetitions during scanning on a scale ranging from 1 (not at all) to 5 (completely). As in Experiment 1, they reported that the memories were not changed very much by repetition (mean: 2.56, SD: 0.81). They were also asked how often they thought of the experience of recalling the memories in Experiment 1 while performing the scanning

task in Experiment 2 on a scale of 1 (not at all) to 5 (during every memory). Participants indicated they rarely thought about Experiment 1 (mean: 1.75, SD: 0.93). Finally, the consistency of recall across time periods during the scanning session did not differ in Experiment 2 ($F_{(7,105)}$ = 0.59, p = 0.761) or between the two experiments ($F_{(1,15)}$ = 0.12, p = 0.733; see also Supplemental Tables 1A and 2A).

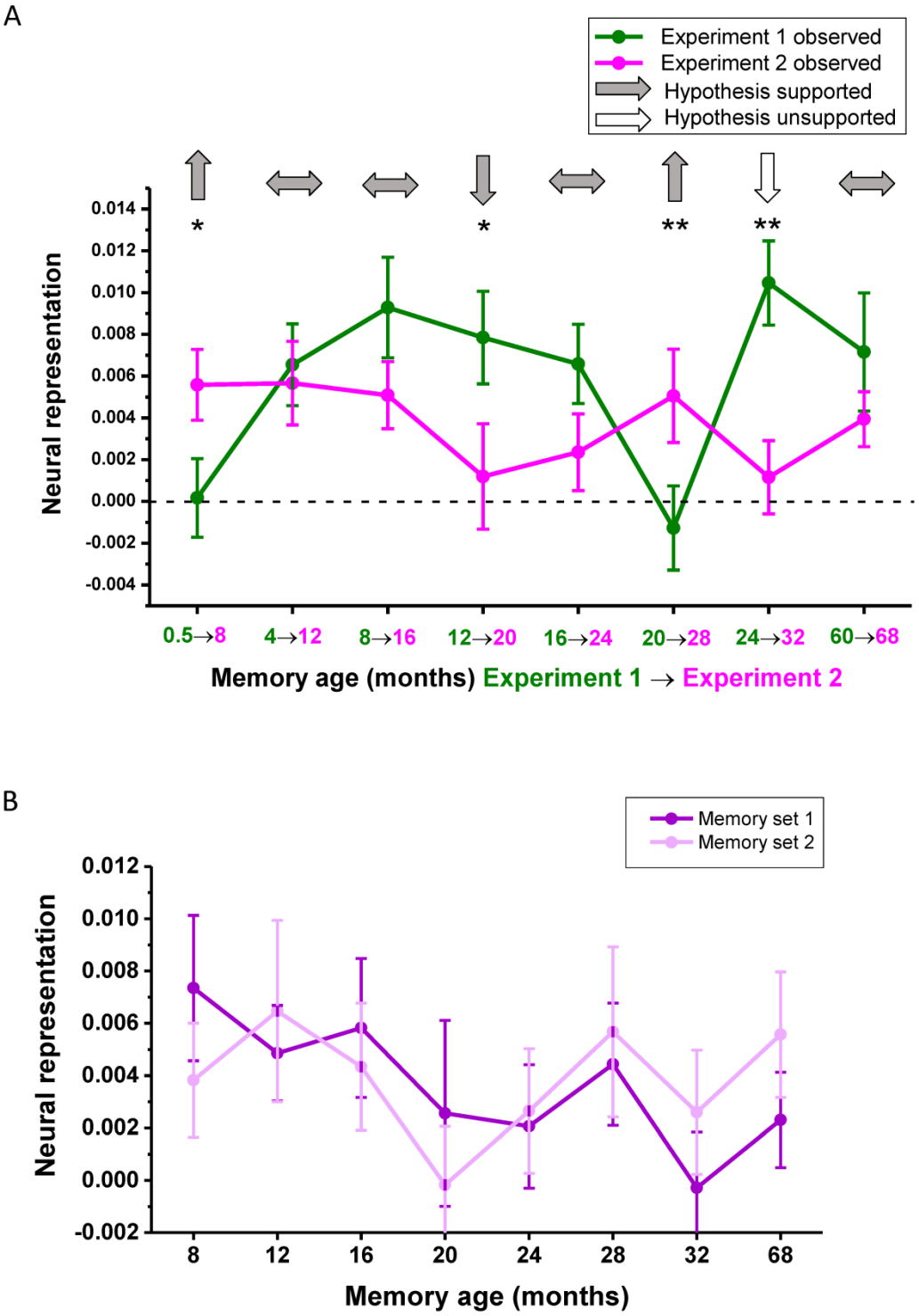

Figure 8. fMRI results of Experiment 2. (**A**) Mean +/- 1SEM neural representation scores at each time-point averaged across the two sets of memories for Experiment 2 (pink line, n=16) compared to the same memories from eight months previously (green line, n=16). Light grey and white arrows indicate supported and unsupported hypotheses respectively; * $p < 0.05$, ** $p < 0.01$. (**B**) Neural representation scores at each time-point for Experiment 2, plotted separately for the two sets of autobiographical memories.

Discussion

This study exploited the sensitivity of RSA to detect not only the extent to which memories of different ages were represented in the vmPFC, but how these representations changed over time. During Experiment 1, we observed detectability in vmPFC for memories at 4M to 12M of age, which was also evident at 24M and 60M. As expected, recent 0.5M old memories were poorly represented in vmPFC in comparison. Curiously, however, the same was observed for memories that were 16M to 20M old. This pattern persisted across separate sets of memories and was replicated in a follow-up study eight months later with the same participants and memories. Behavioural data failed to account for these time-dependent representational changes in either experiment, and this pattern was not evident in other brain areas that we examined.

Of the extant theoretical frameworks of long-term memory consolidation (e.g., Maguire & Mullally, 2013; Lynn Nadel, Winocur, Ryan, & Moscovitch, 2007; Squire & Alvarez, 1995; Teyler & DiScenna, 1985; Winocur & Moscovitch, 2011), these findings speak to none of them. While disagreement persists over the sufficiency of neocortical recruitment for remote autobiographical memory recall, the linear trajectory of systems-level consolidation to the neocortex has never been called into question. There is a growing acceptance of memory consolidation as a fluid and open-ended process (Dudai, 2012). However, the biphasic time-course of engagement observed here prompts a substantial re-evaluation of the vmPFC's contribution to systems-level consolidation.

The lack of vmPFC engagement for the most recent memories is consistent with a time-dependent recruitment. This likely reflects the fact that offline consolidation is in its early stages and a stable memory representation has not yet formed in vmPFC. The nature of this emerging memory trace could be informed by the increasing evidence of a role for vmPFC in the formation and use of schema. This refers to the abstraction of elements common to multiple experiences which help guide future memory recall by constraining the search to representations matching that template

(Hebscher & Gilboa, 2016). Recent memories may not rely on this process to such a large extent, either because they have not yet been assimilated into existing schemas, or are sufficiently close to the current spatiotemporal context to not require such reorientation to representations of the past.

For memories that were 4M to 12M of age we observed a progressive increase in memory detectability in vmPFC. This could reflect the increased adoption of relevant schema to retrieve a memory, as the vmPFC integrates established memories (Schlichting & Preston, 2015). But with integration comes interference. The more fused and embedded within other memories a single representation becomes, the more difficult it is to avoid drifting into connected memories during recall. The resistance of patients with vmPFC damage to the lure of schematically related content during retrieval highlights this natural propensity in healthy controls (Warren, Jones, Duff, & Tranel, 2014). Given that the most recent 12 months represent a congested memory space (Crovitz & Schiffman, 1974), retrieval is likely to also depend on another proposed function of the vmPFC in memory - suppressing those memories which are not relevant (Eichenbaum, 2017). Patients with vmPFC lesions tend to confuse memories from different events (Schnider, von Daniken, & Gutbrod, 1996). This has been attributed to a preconscious filtering out of irrelevant traces, a process which when impaired leads to spontaneous confabulation in patients (Schnider, 2003). Therefore, memory retrieval during this period may represent a delicate balance between locating a memory through the elements it shares with others, and then reliving it through suppressing them.

Across both experiments, retrieval-related neural signatures in vmPFC weaken from 12M to 20M. This suggests they share something in common with the most recent memories, that the aforementioned processes are not critical in their retrieval during this time. Why would this be? Solving this conundrum may entail an appreciation of the flipside of consolidation, that of forgetting. Forgetting is a necessity for a memory system to function optimally (Hardt, Nader, & Nadel, 2013). Therefore, consolidation involves both the strengthening of some memory traces over time, and the

weakening of others (Rauchs et al., 2011), possibly due to their anticipated future relevance (Wamsley, Hamilton, Graveline, Manceor, & Parr, 2016). The forgetting curve of autobiographical memory is particularly steep over the first year (Rubin, 1982). Therefore, the months that follow may represent the point at which a single memory trace has reached optimal stability through consolidation, with minimal interference from previously related events which have now decayed. As a result neither the guidance of a relevant schema, nor inhibition of irrelevant memories are essential for its recall at this time.

It is reasonable to assume that even the most resilient of memory traces eventually succumb to the passage of time. Accordingly, with the most remote memories of 24M and 60M we observed a robust re-engagement of the vmPFC. This could possibly involve the re-instantiation of old schemas to cue fading memory traces which share less and less spatiotemporal features of the present with the passage of time. To clarify, the acceptance of memory trace decay does not necessitate a corresponding decay in memory quality. The comparable amount of details recalled across time periods argue against that supposition. Rather it requires a more distributed network of brain activity (Westphal, Wang, & Rissman, 2017) to compensate for a weaker trace and maintain the consistency of recall over time, which in this case amounts to more engagement of the vmPFC. Furthermore, schema-assisted recruitment does not imply the re-experienced memory is of a schematic or impoverished nature (Winocur & Moscovitch, 2011). The objective scores of internal (episodic) and external (semantic) details were well-matched within and across experiments. So the proposed use of schema here would represent a rapid, preconscious confinement of the memory search to a subset of temporally distant representations.

These interpretations are of course speculative, but also presumptive - that the neural patterns represent content-related processes rather than either the content or process alone. The neural patterns could theoretically represent memory-specific content, but this is not easily reconciled with

the observed biphasic time-course. Conversely, if the patterns represented a generic process common to all autobiographical memory retrieval, detection of individual memories would be impossible as their patterns would not be sufficiently unique. Therefore, it is more reasonable to assume that, when required, the vmPFC retrieves and processes the content of each individual memory in a consistent fashion. That is not to say that structural changes in the region are unnecessary for recall, in fact they may be a prerequisite (Bero et al., 2014; Restivo et al., 2009; Vetere et al., 2011). Rather, the maturing and incorporation of prefrontal cells into a representation over time (Kitamura et al., 2017) allows this region to strategically and flexibly retrieve or reject memories based on their commonalities (Morrissey, Insel, & Takehara-Nishiuchi, 2017) and differences (Guise & Shapiro, 2017). The current results suggest that the relationship between such employed strategies and the passage of time is not linear.

If one is to assign more weight to subjective ratings of detail over objective scores, an alternative explanation for higher neural representation scores at remote time periods is possible. Perhaps older memories are recalled with greater consistency because they are relatively impoverished compared to recent memories. A reduction in available details could result in less recall variability across trials and increased within-memory similarity. Conversely, the richer detail of recent memories would result in a slightly different emphasis during each trial and lower within-memory similarity. If anything we observed the opposite, the only difference observed here was that poorly represented 0.5M old memories were recalled more consistently than 60 month old memories during Experiment 1. A further point on consistency is worth considering, albeit across experiments. It has been suggested that the recall of remote memories could involve the insertion of new episodic details or an increased amount of semantic details, accounting for their differential neural patterns from recent memories (Berkers & van Kesteren, 2013). Here, an inspection of interview transcripts from both experiments revealed participants rarely offered new details upon their return for

Experiment 2, and the analysis of objective scoring of memory details showed no differences between experiments.

Aside from consistency across repetitions, an additional concern in an experiment of this nature is the influence of repetition itself. Memory recall initiates reconsolidation, a transient labile state whereby memory traces become as vulnerable to interference as the time shortly after their initial inception (Nader, 2015). Animal research suggests this lability arises from a re-engagement of the cellular machinery which facilitated consolidation in the first place (Nader, Schafe, & Le Doux, 2000), and possibly serves to strengthen the original memory trace (Lee, 2008). Human episodic memory is also sensitive to such disruption (Hupbach, Gomez, Hardt, & Nadel, 2007) or strengthening during recall (Forcato, Rodriguez, & Pedreira, 2011), while autobiographical memories are not immune to the effects of reconsolidation (Schwabe & Wolf, 2009). Moreover, repeated recall of autobiographical memories has been shown to increase activity of the vmPFC (Nadel, Campbell, & Ryan, 2007). Therefore, retrieval during experimentation could theoretically perturb the natural trajectory of consolidation processes. Verbally recalling the memories one week prior to the initial scan is the first time-window of theoretical interference. However, eliciting memory retrieval one week before the scan effectively resets the recall recency of all memories removing it as a confound. Furthermore, if recall one week before had somehow accelerated the consolidation process, one would expect 0.5M old memories to be more detectable in the vmPFC than they were. The second and third potential reconsolidation windows were the repeated mental recall in the scanner during Experiment 1, and overt memory recall during the Experiment 2 interview. These could theoretically alter the neural data of Experiment 2. However, given that seven out of the eight specifically hypothesised temporally sensitive changes in neural representations were supported, an altered consolidation time-course appears highly unlikely. Again, the recency of memory recall was now matched for Experiment 2, and participants reported very low frequency of recall between experiments. This suggests that repeatedly recalling the memories during the first experiment did

not affect the rate at which participants recalled them subsequently. Increased representational similarity across repetitions also predicts subsequent retrieval success (Xue et al., 2010). Despite the differences in representational similarity scores within Experiment 1, this did not appear to exert an influence as we did not observe any significant change in the number of details recalled between the two experiments.

One other possible interpretation of the unexpected engagement and disengagement of the vmPFC for memories of different ages is that it may be mirrored by a systematic change in the content of memories. For example, types of events that have taken place at a particular time of year which may be common to all participants, such as a seasonal holiday. However, participants were recruited over a period of five months in an evenly spaced manner, making it unlikely that such events would fall into the same temporal windows across participants. The occurrence of personal events such as birthdays would also be naturally random across participants. The use of personal photographs as memory cues also limited the reliance on time of year as a method for strategically retrieving memories. Furthermore, the nature of memory sampling was that unique, rather than generic, events were eligible, reducing the likelihood of events which are repeated annually being included. Memory detectability was high at 12 month intervals such as one, two and five years in this study, suggesting perhaps it is easier to recall events which have taken place at a similar time of year to the present. However this should have been reflected in behavioural ratings, and equivalently strong neural representations for recent memories, but neither was observed. Most importantly, if content rather than time-related consolidation was the main influence on memory detectability, then we would not have observed any change in neural representation scores from Experiment 1 to Experiment 2, rather than the hypothesised shifts which emerged.

In the light of our hypotheses, Experiment 2 generated one anomalous finding. Twenty-four month old memories from Experiment 1 were no longer well represented eight months later. If the

interpretation of the results of Experiment 1 is correct, these memories were originally the most challenging to retrieve, requiring intervention of the vmPFC. In this context, the lower neural representation values for Experiment 2 imply they became less challenging to recall. So perhaps these particular memories were disproportionally affected by a reconsolidation process whereby the repeated vivid recollection in Experiment 1 strengthened the memory trace and reduced the reliance on the vmPFC for Experiment 2. An additional possibility is that memories of around 32M of age are simply not as reliant on vmPFC, and that the biphasic pattern we observed is in fact a feature that iterates again between 24M and 60M. We cannot verify this in the current experiment, as we did not sample this time-period during Experiment 1.

A methodological discrepancy between this experiment and that conducted by Bonnici et al. (2012), is the additional use of a photograph to assist in cueing memories. One possible interpretation of the neural representation scores is they represent a role for the vmPFC in the maintenance of visual working memory following cue offset. However, the prefrontal cortex is unlikely to contribute to maintenance of visual information (Lee, Kravitz, & Baker, 2013). Furthermore, given that we have differing time periods across both experiments, associated with different photographic cues, where memories are undetectable in the vmPFC, this is an inadequate explanation. There is, however, an obvious inconsistency between the findings of the current study and that of Bonnici et al. (2012). Unlike that study, we did not detect representations of 0.5M old memories in vmPFC. It could be that classification-based MVPA is more sensitive to detection of memory representations than RSA, however, the current study was not optimised for such an analysis because it necessitated an increased ratio of conditions to trials. The Bonnici studies also involved many fewer memories that were recalled more times, which may also have also influenced their results. Nonetheless, the increase in memory representation scores from recent to remote memories been replicated and additionally refined in the current study with superior temporal precision.

Given that the medial prefrontal cortex is often associated with value and emotional processing (Grabenhorst & Rolls, 2011), could these factors have influenced the current findings? Humans display a bias towards consolidating positive memories (Anderson & Hanslmayr, 2014) and remembered information is more likely to be valued than that which is forgotten (Rhodes, Witherby, Castel, & Murayama, 2017). Activity in the vmPFC during autobiographical memory recall has been found to be modulated by both the personal significance and emotional content of memory (Lin, Horner, & Burgess, 2016). However, in the current two experiments memories were matched across time periods on these variables. In the eight months between experiments, some memories actually decreased slightly in their subjective ratings of significance and positivity, suggesting that these factors are an unlikely driving force behind the observed remote memory representations in the vmPFC.

In conclusion, the current results revealed a two-stage systems-level consolidation process which was remarkably preserved across completely different sets of memories in one experiment, and closely replicated in a subsequent longitudinal experiment with the same participants and memories. They support the notion that the vmPFC becomes increasingly important over time for the retrieval of remote memories, perhaps by indexing and processing memory traces elsewhere in the neocortex. Two particularly novel findings emerged. First, this process occurs relatively quickly, by four months following an experience. Second, vmPFC involvement after this time fluctuates in a highly consistent manner, depending on the precise age of the memory in question. Further work is clearly needed to explore the implications of these novel findings, including studies looking at vmPFC connectivity with other brain areas such as the hippocampus. Overall, we conclude that our vmPFC findings may be explained by a dynamic interaction between the changing strength of a memory trace, the availability of temporally adjacent memories, and the resultant differential neural circuitry recruited to successfully retrieve the past. The path to consolidation may not be long, but it is winding.

Materials and Methods

Experiment 1

Participants

Thirty healthy, right handed participants (23 female) took part (mean age 25.3, SD 3.5, range 21-32). All had normal or corrected-to-normal vision. Each participant gave written informed consent for participation in the study, for data analysis and for publication of the study results. Materials and methods were approved by the University College London Research Ethics Committee.

Task and procedure

Memory interview and selection of autobiographical memories

Participants were instructed to select at least three photographs from each of eight time-points in their past (0.5M, 4M, 8M, 12M, 16M, 20M, 24M and 60M relative to the time of taking part in the experiment) which reminded them of vivid, unique and specific autobiographical events. Highly personal, emotionally negative or repetitive events were deemed unsuitable. An additional requirement was that memories from the same time period should be dissimilar in content. For the four most recent time periods (0.5M-12M), the memories should have taken place within a temporal window two weeks either side of the specified date, for the next three time points (16M-24M), three weeks either side, and one month either side for the most remote time point (60M). This graded approach was adopted to balance temporal precision with the availability of suitable memories at more remote time-points.

Participants were asked to describe in as much detail as possible the specific autobiographical memory elicited by a photograph. General probes were given by the interviewer where appropriate (e.g., “what else can you remember about this event?”). Participants were also asked to identify the most memorable part of the event which took place within a narrow temporal window and unfolded

in an event-like way. They then created a short phrase pertaining to this episode, which was paired with the photograph to facilitate recall during the subsequent fMRI scan (see Figure 1A). Participants were asked to rate each memory on a number of characteristics (see main text, Figures 1 and 6, Supplemental Tables 1 and 2), and two memories from each time period which satisfied the criteria of high vividness and detail, and ease of recall were selected for recollection during the fMRI scan.

Behavioural Analyses

The interview was recorded and transcribed to facilitate an objective analysis of the details, and the widely-used Autobiographical Interview method was employed for scoring (Levine et al., 2002). Details provided for each memory were scored as either "internal" (specific events, temporal references, places, perceptual observations and thoughts or emotions) or "external" (unrelated events, semantic knowledge, repetition of details or other more general statements). To assess inter-rater reliability, a subset of sixteen memories (n=2 per time period) were randomly selected across 16 different subjects and scored by another experimenter blind to the aims and conditions of the study. Intra-class coefficient estimates were calculated using SPSS statistical package version 22 (SPSS Inc, Chicago, IL) based on a single measures, absolute-agreement, 2-way random-effects model.

As two memories per time period were selected for later recall in the scanner, behavioural ratings were averaged to produce one score per time period. Differences in subjective memory ratings across time periods were analysed using a one-way repeated measures ANOVA with Bonferroni-corrected paired t-tests. Differences in objective memory scores of internal and external details across time periods were analysed using a two-way repeated measures ANOVA with Bonferroni-corrected paired t-tests. A threshold of $p < 0.05$ was used throughout both experiments. All ANOVAs were subjected to Greenhouse-Geisser adjustment to the degrees of freedom if Mauchly's sphericity test identified that sphericity had been violated.

Task during fMRI scanning

Participants returned approximately one week later (mean 6.9 days, SD 1) to recall the memories while undergoing an fMRI scan. Prior to the scan, participants were trained to recall each of the 16 memories within a 12 second recall period (as in Bonnici et al., 2012 and Bonnici and Maguire, 2017), when cued by the photograph alongside its associated cue phrase. There were two training trials per memory, and participants were asked to vividly and consistently recall a particular period of the original event which unfolded across a temporal window matching the recall period.

During scanning, participants recalled each memory six times (6 trials x 16 memories = 96 trials). The two memories from each time period were never recalled together in the same session, nor was any one memory repeated within each session, resulting in 12 separate short sessions with eight trials in each, an approach recommended for optimal detection of condition-related activity patterns using MVPA (Coutanche & Thompson-Schill, 2012). Trials were presented in a random order within each session. On each trial, the photograph and associated pre-selected cue phrase relating to each event were displayed on screen for three seconds. Following removal of this cue, participants then closed their eyes and recalled the memory. After 12 seconds, the black screen flashed white twice, to cue the participant to open their eyes. The participant was then asked to rate how vivid the memory recall had been using a five-key button box, on a scale of 1-5, where 1 was not vivid at all, and 5 was highly vivid. When the least vivid trials were excluded, the mean number of trials (/6) selected for analysis from each time-point were as follows: 0.5M: 5.65 (SD 0.57), 4M: 5.50 (SD 0.56), 8M: 5.43 (SD 0.55), 12M: 5.50 (SD 0.63), 16M: 5.50 (SD 0.59), 20M: 5.43 (SD 0.65), 24M: 5.42 (SD 0.56), 60M: 5.23 (SD 0.69).

MRI data acquisition

Structural and functional data were acquired using a 3T Siemens Trio MRI system (Siemens, Erlangen, Germany). Both types of scan were performed within a partial volume which incorporated the entire extent of the ventromedial prefrontal cortex (see Figure 3A).

Structural images were collected using a single-slab 3D T2-weighted turbo spin echo sequence with variable flip angles (SPACE) (Mugler et al., 2000) in combination with parallel imaging, to simultaneously achieve a high image resolution of ~500 µm, high sampling efficiency and short scan time while maintaining a sufficient signal-to-noise ratio (SNR). After excitation of a single axial slab the image was read out with the following parameters: resolution = 0.52 x 0.52 x 0.5 mm, matrix = 384 x 328, partitions = 104, partition thickness = 0.5 mm, partition oversampling = 15.4%, field of view = 200 x 171 mm 2, TE = 353 ms, TR = 3200 ms, GRAPPA x 2 in phase-encoding (PE) direction, bandwidth = 434 Hz/pixel, echo spacing = 4.98 ms, turbo factor in PE direction = 177, echo train duration = 881, averages = 1.9. For reduction of signal bias due to, for example, spatial variation in coil sensitivity profiles, the images were normalized using a prescan, and a weak intensity filter was applied as implemented by the scanner's manufacturer. To improve the SNR of the anatomical image, three scans were acquired for each participant, coregistered and averaged. Additionally, a whole brain 3D FLASH structural scan was acquired with a resolution of 1 x 1 x 1 mm.

Functional data were acquired using a 3D echo planar imaging (EPI) sequence which has been demonstrated to yield improved BOLD sensitivity compared to 2D EPI acquisitions (Lutti, Thomas, Hutton, & Weiskopf, 2013). Image resolution was $1.5mm^3$ and the field-of-view was 192mm in-plane. Forty slices were acquired with 20% oversampling to avoid wrap-around artefacts due to imperfect slab excitation profile. The echo time (TE) was 37.30 ms and the volume repetition time (TR) was 3.65s. Parallel imaging with GRAPPA image reconstruction (Griswold et al., 2002) acceleration factor 2 along the phase-encoding direction was used to minimize image distortions and

yield optimal BOLD sensitivity. The dummy volumes necessary to reach steady state and the GRAPPA reconstruction kernel were acquired prior to the acquisition of the image data as described in Lutti et al. (2013). Correction of the distortions in the EPI images was implemented using B0-field maps obtained from double-echo FLASH acquisitions (matrix size 64x64; 64 slices; spatial resolution $3mm^3$; short TE=10 ms; long TE=12.46 ms; TR=1020 ms) and processed using the FieldMap toolbox available in SPM (Hutton et al., 2002).

MRI data analysis

Preprocessing

fMRI data were analysed using SPM12 (www.fil.ion.ucl.ac.uk/spm). All images were first bias corrected to compensate for image inhomogeneity associated with the 32 channel head coil (Van Leemput, Maes, Vandermeulen, & Suetens, 1999). Fieldmaps collected during the scan were used to generate voxel displacement maps. EPIs for each of the twelve sessions were then realigned to the first image and unwarped using the voxel displacement maps calculated above. The three high-resolution structural images were averaged to reduce noise, and co-registered to the whole brain structural scan. EPIs were also co-registered to the whole brain structural scan. Manual segmentation of the vmPFC was performed using ITK-SNAP on the group averaged structural scan normalised to MNI space. The normalised group mask was warped back into each participant's native space using the inverse deformation field generated by individual participant structural scan segmentations. The overlapping voxels between this participant-specific vmPFC mask and the grey matter mask generated by the structural scan segmentation were used to create a native-space grey matter vmPFC mask for each individual participant.

Representational Similarity Analysis

Functional data were analysed at the single subject level without warping or smoothing. Each recall trial was modelled as a separate GLM, which comprised the 12 second period from the offset of the

memory cue to just before the white flash which indicated to the participant they should open their eyes. Motion parameters were included as regressors of no interest. RSA (Kriegeskorte & Kievit, 2013), was performed using the RSA toolbox (http://www.mrc-cbu.cam.ac.uk/methods-and-resources/toolboxes/) and custom MATLAB scripts (version R2014a). In order to account for the varying levels of noise across voxels which can affect the results of multivariate fMRI analyses, multivariate noise normalisation (Walther et al., 2016) was performed on the estimated pattern of neural activity separately for each trial. This approach normalises the estimated beta weight of each voxel using the residuals of the first-level GLM and the covariance structure of this noise. This results in the down-weighting of noisier voxels and a more accurate estimate of the task-related activity of each voxel.

The average number of voxels analysed in the vmPFC across the two sets of memories was 5252 (SD: 1227). Whole ROI-based analysis was preferred to a searchlight approach which would involve comparing neural with model similarity matrices (Kriegeskorte, Goebel, & Bandettini, 2006), as we did not have very strong *a priori* hypotheses about changes in neural representations over time against which to test the neural data, nor did we want to make assumptions regarding the spatial distribution of informative voxels in the vmPFC.

As participants recalled two memories per time-point, the dataset was first split into two sets of eight time points, which were analysed separately using RSA. To characterise the strength of memory representations in the vmPFC, the similarity of neural patterns across recall trials of the same memory was first calculated using the Pearson product-moment correlation coefficient, resulting in a “within-memory” similarity score. Then the neural patterns of each memory were correlated with those of all other memories, yielding a “between-memory” similarity score. For each memory, the between-memory score was then subtracted from the within-memory score to provide a neural representation score (see Figure 3C). This score was then averaged across the two

memories at each time-point. Results for the left and the right hemispheres were highly similar, and therefore the data we report here are from the vmPFC bilaterally. A distinctive neural pattern associated with the recall of memories at each time period would yield a score significantly higher than zero, which was assessed using a one-sample t-test. Strengthening or weakening of memory representations as a function of remoteness would result in a significant difference in memory representation scores across time periods, and this was assessed using a one-way repeated measures ANOVA with post-hoc two-tailed paired t-tests. The range of values that we observed are entirely consistent with those in other studies employing a similar RSA approach in a variety of learning, memory and navigation tasks in a wide range of brain regions (Bellmund, Deuker, Navarro Schröder, & Doeller, 2016; Chadwick, Jolly, Amos, Hassabis, & Spiers, 2015; Deuker, Bellmund, Navarro Schröder, & Doeller, 2016; Hsieh, Gruber, Jenkins, & Ranganath, 2014; Hsieh & Ranganath, 2015; Kim, Jeffery, & Maguire, 2017; Milivojevic, Vicente-Grabovetsky, & Doeller, 2015; Schapiro, Turk-Browne, Norman, & Botvinick, 2016; Schuck, Cai, Wilson, & Niv, 2016; Staresina, Henson, Kriegeskorte, & Alink, 2012).

Of note, we also conducted a standard mass-univariate analysis on the data with memory remoteness as a parametric regressor, and this did not reveal any significant results, consistent with the findings of Bonnici et al. (2012).

Experiment 2

Participants

Sixteen of the 30 participants who took part in Experiment 1 returned to take part in Experiment 2 (14 female, mean age 24.7, SD 3.1, range 21-33) approximately eight months later (8.4 months, SD 1.2).

Task and procedure

Memory interview

Participants were presented with the 16 photographs and cue phrases associated with the autobiographical memories in Experiment 1 and were asked to describe in as much detail as possible the specific event which they had recalled previously. General probes were given by the interviewer where appropriate (e.g. "what else can you remember about this event?"). The interviewer availed of summarised transcripts from Experiment 1 to verify the same memory and details were being recalled. Participants then rated each memory on the same characteristics assessed in Experiment one. The memory interview during Experiment 2 was also recorded and transcribed.

Behavioural Analyses

The analysis of subjective and objective ratings for Experiment 2 followed exactly the same procedure as Experiment 1. The extent to which subjective ratings for the same memory had changed between Experiment 1 and Experiment 2 was assessed using a two-way (experiment x time period) repeated measures ANOVA with Bonferroni-corrected paired t-tests. Differences in objective memory ratings across experiments were analysed using a two (experiment) x two (detail) x eight (time period) repeated measures ANOVA with Bonferroni-corrected paired t-tests.

Task during fMRI scanning

Participants returned approximately one week later for the fMRI scan (mean 5.5 days, SD 3.7). Prior to scanning, only one reminder training trial per memory was deemed necessary given the prior experience of performing the task in Experiment 1. The scanning task remained unchanged from Experiment 1, aside from the re-randomisation of trials within each session. When the least vivid trials were excluded, the mean number of trials (/6) selected for analysis from each time period were as follows: 8M: 5.94 (SD 0.25), 12M: 5.97 (SD 0.13), 16M: 5.88 (SD 0.29), 20M: 5.88 (SD 0.29), 24M: 5.94 (SD 0.25), 28M: 5.94 (SD 0.17), 32M: 5.84 (SD 0.40), 68M: 5.81 (SD 0.36).

MRI data acquisition

Structural and functional data were acquired using the same scanner and scanning sequences as Experiment 1. However the prior acquisition of the partial volume structural scans negated the need to include these in the protocol of Experiment 2.

MRI data analysis

Preprocessing

fMRI data were preprocessed using the same pipeline as Experiment 1, with the additional step of co-registering the functional scans of Experiment 2 to the structural scans of Experiment 1, which enabled the use of the vmPFC masks from Experiment 1. First-level GLMs of each recall trial were constructed in an identical manner to Experiment 1.

Representational Similarity Analysis

RSA of the Experiment 2 fMRI data was conducted in an identical manner to Experiment 1. The average number of voxels analysed in the vmPFC across the two sets of memories for all participants was 5228 (SD: 1765). To generate predicted changes in representations in the eight months from Experiment 1 to Experiment 2, the scores from Experiment 1 were shifted by two time-points, and a two-tailed paired t-test was performed on each memory's original neural representation score and its expected score eight months later (see Figure 5). To ascertain whether the observed neural representation scores had changed between Experiments 1 and 2, a two-way (experiment x time period) repeated measures ANOVA was performed. To investigate if these changes mirrored the predictions generated by the original data, paired t-tests were performed between the actual neural representation scores for each memory from Experiment 1 and Experiment 2, one-tailed if there was a hypothesised increase or decrease.

Acknowledgements

We thank David Bradbury and Imaging Support for technical assistance.

Ethics

The studies were approved by the University College London Research Ethics Committee: #6743/002 Systems-Level Consolidation of Autobiographical Memories. Written informed consent was obtained from each participant for participation in the study, for data analysis and for publication of the study results.

References

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends Cogn Sci, 18*(6), 279-292. doi:10.1016/j.tics.2014.03.002

Bellmund, J. L. S., Deuker, L., Navarro Schröder, T., & Doeller, C. F. (2016). Grid-cell representations in mental simulation. *eLife, 5*, e17089. doi:10.7554/eLife.17089

Berkers, R. M., & van Kesteren, M. T. (2013). Autobiographical memory transformation across consolidation. *J Neurosci, 33*(13), 5435-5436. doi:10.1523/JNEUROSCI.0246-13.2013

Bero, A. W., Meng, J., Cho, S., Shen, A. H., Canter, R. G., Ericsson, M., & Tsai, L. H. (2014). Early remodeling of the neocortex upon episodic memory encoding. *Proc Natl Acad Sci U S A, 111*(32), 11852-11857. doi:10.1073/pnas.1408378111

Bertossi, E., Tesini, C., Cappelli, A., & Ciaramelli, E. (2016). Ventromedial prefrontal damage causes a pervasive impairment of episodic memory and future thinking. *Neuropsychologia, 90*, 12-24. doi:10.1016/j.neuropsychologia.2016.01.034

Bonnici, H. M., Chadwick, M. J., Lutti, A., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2012). Detecting representations of recent and remote autobiographical memories in vmPFC and hippocampus. *J Neurosci, 32*(47), 16982-16991. doi:10.1523/JNEUROSCI.2475-12.2012

Bonnici, H. M., & Maguire, E. A. (2017). Two years later - Revisiting autobiographical memory representations in vmPFC and hippocampus. *Neuropsychologia*. doi:10.1016/j.neuropsychologia.2017.05.014

Chadwick, M. J., Bonnici, H. M., & Maguire, E. A. (2012). Decoding information in the human hippocampus: a user's guide. *Neuropsychologia, 50*(13), 3107-3121. doi:10.1016/j.neuropsychologia.2012.07.007

Chadwick, M. J., Jolly, A. E. J., Amos, D. P., Hassabis, D., & Spiers, H. J. (2015). A goal direction signal in the human entorhinal/subicular region. *Current Biology, 25*(1), 87-92. doi:http://dx.doi.org/10.1016/j.cub.2014.11.001

Coutanche, M. N., & Thompson-Schill, S. L. (2012). The advantage of brief fMRI acquisition runs for multi-voxel pattern detection across runs. *Neuroimage, 61*(4), 1113-1119. doi:10.1016/j.neuroimage.2012.03.076

Crovitz, H. F., & Schiffman, H. (1974). Frequency of Episodic Memories as a Function of Their Age. *Bulletin of the Psychonomic Society, 4*(Nb5), 517-518.

Deuker, L., Bellmund, J. L. S., Navarro Schröder, T., & Doeller, C. F. (2016). An event map of memory space in the hippocampus. *eLife, 5*, e16534. doi:http://dx.doi.org/10.7554/eLife.16534

Ding, H. K., Teixeira, C. M., & Frankland, P. W. (2008). Inactivation of the anterior cingulate cortex blocks expression of remote, but not recent, conditioned taste aversion memory. *Learn Mem, 15*(5), 290-293. doi:10.1101/lm.905008

Dudai, Y. (2012). The restless engram: consolidations never end. *Annu Rev Neurosci, 35*, 227-247. doi:10.1146/annurev-neuro-062111-150500

Eichenbaum, H. (2017). Memory: Organization and Control. *Annu Rev Psychol, 68*, 19-45. doi:10.1146/annurev-psych-010416-044131

Einarsson, E. O., & Nader, K. (2012). Involvement of the anterior cingulate cortex in formation, consolidation, and reconsolidation of recent and remote contextual fear memory. *Learn Mem, 19*(10), 449-452. doi:10.1101/lm.027227.112

Forcato, C., Rodriguez, M. L., & Pedreira, M. E. (2011). Repeated labilization-reconsolidation processes strengthen declarative memory in humans. *PLoS One, 6*(8), e23305. doi:10.1371/journal.pone.0023305

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nat Rev Neurosci, 6*(2), 119-130. doi:10.1038/nrn1607

Frankland, P. W., Bontempi, B., Talton, L. E., Kaczmarek, L., & Silva, A. J. (2004). The involvement of the anterior cingulate cortex in remote contextual fear memory. *Science, 304*(5672), 881-883. doi:10.1126/science.1094804

Furman, O., Mendelsohn, A., & Dudai, Y. (2012). The episodic engram transformed: Time reduces retrieval-related brain activity but correlates it with memory accuracy. *Learn Mem, 19*(12), 575-587. doi:10.1101/lm.025965.112

Gilboa, A., Winocur, G., Grady, C. L., Hevenor, S. J., & Moscovitch, M. (2004). Remembering our past: functional neuroanatomy of recollection of recent and very remote personal events. *Cereb Cortex, 14*(11), 1214-1225. doi:10.1093/cercor/bhh082

Grabenhorst, F., & Rolls, E. T. (2011). Value, pleasure and choice in the ventral prefrontal cortex. *Trends Cogn Sci, 15*(2), 56-67. doi:10.1016/j.tics.2010.12.004

Griswold, M. A., Jakob, P. M., Heidemann, R. M., Nittka, M., Jellus, V., Wang, J., . . . Haase, A. (2002). Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med, 47*(6), 1202-1210. doi:10.1002/mrm.10171

Guise, K. G., & Shapiro, M. L. (2017). Medial Prefrontal Cortex Reduces Memory Interference by Modifying Hippocampal Encoding. *Neuron, 94*(1), 183-192 e188. doi:10.1016/j.neuron.2017.03.011

Guzowski, J. F. (2002). Insights into immediate-early gene function in hippocampal memory consolidation using antisense oligonucleotide and fluorescent imaging approaches. *Hippocampus, 12*(1), 86-104. doi:10.1002/hipo.10010

Harand, C., Bertran, F., La Joie, R., Landeau, B., Mezenge, F., Desgranges, B., . . . Rauchs, G. (2012). The hippocampus remains activated over the long term for the retrieval of truly episodic memories. *PLoS One, 7*(8), e43495. doi:10.1371/journal.pone.0043495

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: the role of active forgetting in memory. *Trends Cogn Sci, 17*(3), 111-120. doi:10.1016/j.tics.2013.01.001

Hebscher, M., & Gilboa, A. (2016). A boost of confidence: The role of the ventromedial prefrontal cortex in memory, decision-making, and schemas. *Neuropsychologia, 90*, 46-58. doi:10.1016/j.neuropsychologia.2016.05.003

Hsieh, L.-T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron, 81*(5), 1165-1178. doi:http://dx.doi.org/10.1016/j.neuron.2014.01.015

Hsieh, L.-T., & Ranganath, C. (2015). Cortical and subcortical contributions to sequence retrieval: Schematic coding of temporal context in the neocortical recollection network. *Neuroimage, 121*, 78-90. doi:http://dx.doi.org/10.1016/j.neuroimage.2015.07.040

Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: a subtle reminder triggers integration of new information. *Learn Mem, 14*(1-2), 47-53. doi:10.1101/lm.365707

Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., & Turner, R. (2002). Image distortion correction in fMRI: A quantitative evaluation. *Neuroimage, 16*(1), 217-240. doi:10.1006/nimg.2001.1054

Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science, 294*(5544), 1030-1038. doi:10.1126/science.1067020

Kim, M., Jeffery, K. J., & Maguire, E. A. (2017). Multivoxel pattern analysis reveals 3D place information in the human hippocampus. *The Journal of Neuroscience*. doi:10.1523/jneurosci.2703-16.2017

Kitamura, T., Ogawa, S. K., Roy, D. S., Okuyama, T., Morrissey, M. D., Smith, L. M., . . . Tonegawa, S. (2017). Engrams and circuits crucial for systems consolidation of a memory. *Science, 356*(6333), 73-78. doi:10.1126/science.aam6808

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proc Natl Acad Sci U S A, 103*(10), 3863-3868. doi:10.1073/pnas.0600244103

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn Sci, 17*(8), 401-412. doi:10.1016/j.tics.2013.06.007

Lee, J. L. (2008). Memory reconsolidation mediates the strengthening of memories by additional learning. *Nat Neurosci, 11*(11), 1264-1266. doi:10.1038/nn.2205

Lee, S. H., Kravitz, D. J., & Baker, C. I. (2013). Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nat Neurosci, 16*(8), 997-999. doi:10.1038/nn.3452

Leon, W. C., Bruno, M. A., Allard, S., Nader, K., & Cuello, A. C. (2010). Engagement of the PFC in consolidation and recall of recent spatial memory. *Learn Mem, 17*(6), 297-305. doi:10.1101/lm.1804410

Lesburgueres, E., Gobbo, O. L., Alaux-Cantin, S., Hambucken, A., Trifilieff, P., & Bontempi, B. (2011). Early tagging of cortical networks is required for the formation of enduring associative memory. *Science, 331*(6019), 924-928. doi:10.1126/science.1196164

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging, 17*(4), 677-689. doi:10.1037//0882-7974.17.4.677

Lin, W. J., Horner, A. J., & Burgess, N. (2016). Ventromedial prefrontal cortex, adding value to autobiographical memories. *Sci Rep, 6*, 28630. doi:10.1038/srep28630

Liu, F., Zheng, X. L., & Li, B. M. (2009). The anterior cingulate cortex is involved in retrieval of long-term/long-lasting but not short-term memory for step-through inhibitory avoidance in rats. *Neurosci Lett, 460*(2), 175-179. doi:10.1016/j.neulet.2009.05.032

Lopez, J., Herbeaux, K., Cosquer, B., Engeln, M., Muller, C., Lazarus, C., . . . de Vasconcelos, A. P. (2012). Context-dependent modulation of hippocampal and cortical recruitment during remote spatial memory retrieval. *Hippocampus, 22*(4), 827-841. doi:10.1002/hipo.20943

Luo, F., Zheng, J., Sun, X., Deng, W. K., Li, B. M., & Liu, F. (2017). Prelimbic cortex extracellular signal-regulated kinase 1/2 activation is required for memory retrieval of long-term inhibitory avoidance. *Brain Res, 1661*, 88-99. doi:10.1016/j.brainres.2017.02.010

Lutti, A., Thomas, D. L., Hutton, C., & Weiskopf, N. (2013). High-resolution functional MRI at 3 T: 3D/2D echo-planar imaging with optimized physiological noise correction. *Magn Reson Med, 69*(6), 1657-1664. doi:10.1002/mrm.24398

Mackey, S., & Petrides, M. (2014). Architecture and morphology of the human ventromedial prefrontal cortex. *Eur J Neurosci, 40*(5), 2777-2796. doi:10.1111/ejn.12654

Maguire, E. A. (2014). Memory consolidation in humans: new evidence and opportunities. *Exp Physiol, 99*(3), 471-486. doi:10.1113/expphysiol.2013.072157

Maguire, E. A., & Frith, C. D. (2003). Lateral asymmetry in the hippocampal response to the remoteness of autobiographical memories. *J Neurosci, 23*(12), 5302-5307.

Maguire, E. A., Henson, R. N., Mummery, C. J., & Frith, C. D. (2001). Activity in prefrontal cortex, not hippocampus, varies parametrically with the increasing remoteness of memories. *Neuroreport, 12*(3), 441-444.

Maguire, E. A., & Mullally, S. L. (2013). The hippocampus: a manifesto for change. *J Exp Psychol Gen, 142*(4), 1180-1189. doi:10.1037/a0033650

Maviel, T., Durkin, T. P., Menzaghi, F., & Bontempi, B. (2004). Sites of neocortical reorganization critical for remote spatial memory. *Science, 305*(5680), 96-99. doi:10.1126/science.1098180

McCormick, C., Ciaramelli, E., De Luca, F., & Maguire, E. A. (2017). Comparing and contrasting the cognitive effects of hippocampal and ventromedial prefrontal cortex damage: A review of human lesion studies. *Neuroscience*. doi:http://dx.doi.org/10.1016/j.neuroscience.2017.07.066

Milivojevic, B., Vicente-Grabovetsky, A., & Doeller, C. F. (2015). Insight Reconfigures Hippocampal-Prefrontal Memories. *Current Biology, 25*(7), 821-830. doi:http://dx.doi.org/10.1016/j.cub.2015.01.033

Morris, R. G., Moser, E. I., Riedel, G., Martin, S. J., Sandin, J., Day, M., & O'Carroll, C. (2003). Elements of a neurobiological theory of the hippocampus: the role of activity-dependent synaptic plasticity in memory. *Philos Trans R Soc Lond B Biol Sci, 358*(1432), 773-786. doi:10.1098/rstb.2002.1264

Morrissey, M. D., Insel, N., & Takehara-Nishiuchi, K. (2017). Generalizable knowledge outweighs incidental details in prefrontal ensemble code over time. *Elife, 6*. doi:10.7554/eLife.22177

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic Memory and Beyond: The Hippocampus and Neocortex in Transformation. *Annu Rev Psychol, 67*, 105-134. doi:10.1146/annurev-psych-113011-143733

Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., . . . Nadel, L. (2005). Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory. *J Anat, 207*(1), 35-66. doi:10.1111/j.1469-7580.2005.00421.x

Mugler, J. P., 3rd, Bao, S., Mulkern, R. V., Guttmann, C. R., Robertson, R. L., Jolesz, F. A., & Brookeman, J. R. (2000). Optimized single-slab three-dimensional spin-echo MR imaging of the brain. *Radiology, 216*(3), 891-899. doi:10.1148/radiology.216.3.r00au46891

Nadel, L., Campbell, J., & Ryan, L. (2007). Autobiographical memory retrieval and hippocampal activation as a function of repetition and the passage of time. *Neural Plast, 2007*, 90472. doi:10.1155/2007/90472

Nadel, L., Winocur, G., Ryan, L., & Moscovitch, M. (2007). Systems consolidation and hippocampus: two views. *Debates in Neuroscience, 1*(2-4), 55-66. doi:10.1007/s11559-007-9003-9

Nader, K. (2015). Reconsolidation and the Dynamic Nature of Memory. *Cold Spring Harb Perspect Biol, 7*(10), a021782. doi:10.1101/cshperspect.a021782

Nader, K., Schafe, G. E., & Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature, 406*(6797), 722-726. doi:10.1038/35021052

Nieuwenhuis, I. L., & Takashima, A. (2011). The role of the ventromedial prefrontal cortex in memory consolidation. *Behav Brain Res, 218*(2), 325-334. doi:10.1016/j.bbr.2010.12.009

Niki, K., & Luo, J. (2002). An fMRI study on the time-limited role of the medial temporal lobe in long-term topographical autobiographic memory. *J Cogn Neurosci, 14*(3), 500-507. doi:10.1162/089892902317362010

Oddo, S., Lux, S., Weiss, P. H., Schwab, A., Welzer, H., Markowitsch, H. J., & Fink, G. R. (2010). Specific role of medial prefrontal cortex in retrieving recent autobiographical memories: an fMRI study of young female subjects. *Cortex, 46*(1), 29-39. doi:10.1016/j.cortex.2008.07.003

Pastalkova, E., Serrano, P., Pinkhasova, D., Wallace, E., Fenton, A. A., & Sacktor, T. C. (2006). Storage of spatial information by the maintenance mechanism of LTP. *Science, 313*(5790), 1141-1144. doi:10.1126/science.1128657

Piefke, M., Weiss, P. H., Markowitsch, H. J., & Fink, G. R. (2005). Gender differences in the functional neuroanatomy of emotional episodic autobiographical memory. *Hum Brain Mapp, 24*(4), 313-324. doi:10.1002/hbm.20092

Piolino, P., Giffard-Quillon, G., Desgranges, B., Chetelat, G., Baron, J. C., & Eustache, F. (2004). Re-experiencing old memories via hippocampus: a PET study of autobiographical memory. *Neuroimage, 22*(3), 1371-1383. doi:10.1016/j.neuroimage.2004.02.025

Rauchs, G., Feyers, D., Landeau, B., Bastin, C., Luxen, A., Maquet, P., & Collette, F. (2011). Sleep contributes to the strengthening of some memories over others, depending on hippocampal activity at learning. *J Neurosci, 31*(7), 2563-2568. doi:10.1523/JNEUROSCI.3972-10.2011

Rekkas, P. V., & Constable, R. T. (2005). Evidence that autobiographic memory retrieval does not become independent of the hippocampus: an fMRI study contrasting very recent with remote events. *J Cogn Neurosci, 17*(12), 1950-1961.

Restivo, L., Vetere, G., Bontempi, B., & Ammassari-Teule, M. (2009). The formation of recent and remote memory is associated with time-dependent formation of dendritic spines in the hippocampus and anterior cingulate cortex. *J Neurosci, 29*(25), 8206-8214. doi:10.1523/JNEUROSCI.0966-09.2009

Rhodes, M. G., Witherby, A. E., Castel, A. D., & Murayama, K. (2017). Explaining the forgetting bias effect on value judgments: The influence of memory for a past test. *Mem Cognit, 45*(3), 362-374. doi:10.3758/s13421-016-0674-z

Rubin, D. C. (1982). On the Retention Function for Autobiographical Memory. *Journal of Verbal Learning and Verbal Behavior, 21*(1), 21-38. doi:Doi 10.1016/S0022-5371(82)90423-6

Runyan, J. D., & Dash, P. K. (2005). Inhibition of hippocampal protein synthesis following recall disrupts expression of episodic-like memory in trace conditioning. *Hippocampus, 15*(3), 333-339. doi:10.1002/hipo.20055

Ryan, L., Nadel, L., Keil, K., Putnam, K., Schnyer, D., Trouard, T., & Moscovitch, M. (2001). Hippocampal complex and retrieval of recent and very remote autobiographical memories: evidence from functional magnetic resonance imaging in neurologically intact people. *Hippocampus, 11*(6), 707-714. doi:10.1002/hipo.1086

Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus, 26*(1), 3-8. doi:http://dx.doi.org/10.1002/hipo.22523

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: neural mechanisms and implications for behavior. *Curr Opin Behav Sci, 1*, 1-8. doi:10.1016/j.cobeha.2014.07.005

Schnider, A. (2003). Spontaneous confabulation and the adaptation of thought to ongoing reality. *Nat Rev Neurosci, 4*(8), 662-671. doi:10.1038/nrn1179

Schnider, A., von Daniken, C., & Gutbrod, K. (1996). Disorientation in amnesia. A confusion of memory traces. *Brain, 119 (Pt 5)*, 1627-1632.

Schuck, Nicolas W., Cai, Ming B., Wilson, Robert C., & Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron, 91*(6), 1402-1412. doi:http://dx.doi.org/10.1016/j.neuron.2016.08.019

Schwabe, L., & Wolf, O. T. (2009). New episodic learning interferes with the reconsolidation of autobiographical memories. *PLoS One, 4*(10), e7519. doi:10.1371/journal.pone.0007519

Soderlund, H., Moscovitch, M., Kumar, N., Mandic, M., & Levine, B. (2012). As time goes by: hippocampal connectivity changes with remoteness of autobiographical memory retrieval. *Hippocampus, 22*(4), 670-679. doi:10.1002/hipo.20927

Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr Opin Neurobiol, 5*(2), 169-177.

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harb Perspect Biol, 7*(8), a021766. doi:10.1101/cshperspect.a021766

Staresina, B. P., Henson, R. N. A., Kriegeskorte, N., & Alink, A. (2012). Episodic reinstatement in the medial temporal lobe. *J Neurosci., 32*(50), 18150-18156. doi:10.1523/jneurosci.4156-12.2012

Steinvorth, S., Corkin, S., & Halgren, E. (2006). Ecphory of autobiographical memories: an fMRI study of recent and remote memory retrieval. *Neuroimage, 30*(1), 285-298. doi:10.1016/j.neuroimage.2005.09.025

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: a meta-analysis. *Neuropsychologia, 44*(12), 2189-2208. doi:10.1016/j.neuropsychologia.2006.05.023

Takashima, A., Nieuwenhuis, I. L., Jensen, O., Talamini, L. M., Rijpkema, M., & Fernandez, G. (2009). Shift from hippocampal to neocortical centered retrieval network with consolidation. *J Neurosci, 29*(32), 10087-10093. doi:10.1523/JNEUROSCI.0799-09.2009

Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M. J., . . . Fernandez, G. (2006). Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. *Proc Natl Acad Sci U S A, 103*(3), 756-761. doi:10.1073/pnas.0507774103

Takehara-Nishiuchi, K., Nakao, K., Kawahara, S., Matsuki, N., & Kirino, Y. (2006). Systems consolidation requires postlearning activation of NMDA receptors in the medial prefrontal cortex in trace eyeblink conditioning. *J Neurosci, 26*(19), 5049-5058. doi:10.1523/JNEUROSCI.4381-05.2006

Takehara, K., Kawahara, S., & Kirino, Y. (2003). Time-dependent reorganization of the brain components underlying memory retention in trace eyeblink conditioning. *J Neurosci, 23*(30), 9897-9905.

Teixeira, C. M., Pomedli, S. R., Maei, H. R., Kee, N., & Frankland, P. W. (2006). Involvement of the anterior cingulate cortex in the expression of remote spatial memory. *J Neurosci, 26*(29), 7555-7564. doi:10.1523/JNEUROSCI.1068-06.2006

Teyler, T. J., & DiScenna, P. (1985). The role of hippocampus in memory: a hypothesis. *Neurosci Biobehav Rev, 9*(3), 377-389.

Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., . . . Morris, R. G. (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science, 333*(6044), 891-895. doi:10.1126/science.1205274

Tsukiura, T., Fujii, T., Okuda, J., Ohtake, H., Kawashima, R., Itoh, M., . . . Yamadori, A. (2002). Time-dependent contribution of the hippocampal complex when remembering the past: a PET study. *Neuroreport, 13*(17), 2319-2323. doi:10.1097/01.wnr.0000044989.13025.79

Turner, M. S., Cipolotti, L., Yousry, T. A., & Shallice, T. (2008). Confabulation: damage to a specific inferior medial prefrontal system. *Cortex, 44*(6), 637-648. doi:10.1016/j.cortex.2007.01.002

Van Leemput, K., Maes, F., Vandermeulen, D., & Suetens, P. (1999). Automated model-based tissue classification of MR images of the brain. *IEEE Trans Med Imaging, 18*(10), 897-908. doi:10.1109/42.811270

Vetere, G., Restivo, L., Cole, C. J., Ross, P. J., Ammassari-Teule, M., Josselyn, S. A., & Frankland, P. W. (2011). Spine growth in the anterior cingulate cortex is necessary for the consolidation of contextual fear memory. *Proc Natl Acad Sci U S A, 108*(20), 8456-8460. doi:10.1073/pnas.1016275108

Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of dissimilarity measures for multi-voxel pattern analysis. *Neuroimage, 137*, 188-200. doi:10.1016/j.neuroimage.2015.12.012

Wamsley, E. J., Hamilton, K., Graveline, Y., Manceor, S., & Parr, E. (2016). Test Expectation Enhances Memory Consolidation across Both Sleep and Wake. *PLoS One, 11*(10), e0165141. doi:10.1371/journal.pone.0165141

Warren, D. E., Jones, S. H., Duff, M. C., & Tranel, D. (2014). False recall is reduced by damage to the ventromedial prefrontal cortex: implications for understanding the neural correlates of schematic memory. *J Neurosci, 34*(22), 7677-7682. doi:10.1523/JNEUROSCI.0119-14.2014

Watanabe, T., Kimura, H. M., Hirose, S., Wada, H., Imai, Y., Machida, T., . . . Konishi, S. (2012). Functional dissociation between anterior and posterior temporal cortical regions during retrieval of remote memory. *J Neurosci, 32*(28), 9659-9670. doi:10.1523/JNEUROSCI.5553-11.2012

Westphal, A. J., Wang, S., & Rissman, J. (2017). Episodic Memory Retrieval Benefits from a Less Modular Brain Network Organization. *J Neurosci, 37*(13), 3523-3531. doi:10.1523/JNEUROSCI.2509-16.2017

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *J Int Neuropsychol Soc, 17*(5), 766-780. doi:10.1017/S1355617711000683

Xue, G., Dong, Q., Chen, C., Lu, Z., Mumford, J. A., & Poldrack, R. A. (2010). Greater neural pattern similarity across repetitions is associated with better memory. *Science, 330*(6000), 97-101. doi:10.1126/science.1193125

Zhao, M. G., Toyoda, H., Lee, Y. S., Wu, L. J., Ko, S. W., Zhang, X. H., . . . Zhuo, M. (2005). Roles of NMDA NR2B subtype receptor in prefrontal long-term potentiation and contextual fear memory. *Neuron, 47*(6), 859-872. doi:10.1016/j.neuron.2005.08.014

Supplemental Table 1, related to Figures 1 and 2: Behavioural data (mean, SD) - Experiment 1 (n=30)

A. Subjective ratings	0.5M	4M	8M	12M	16M	20M	24M	60M
Vividness	4.53 (0.52)	4.28 (0.55)	4.22 (0.50)	4.05 (0.53)	4.25 (0.50)	4.10 (0.50)	4.10 (0.52)	4.13 (0.67)
Detail	4.38 (0.54)	3.88 (0.61)	3.93 (0.49)	3.67 (0.75)	3.80 (0.60)	3.68 (0.71)	3.80 (0.57)	3.77 (0.78)
Effort	1.25 (0.39)	1.53 (0.45)	1.52 (0.50)	1.87 (0.57)	1.72 (0.47)	1.58 (0.44)	1.57 (0.49)	1.77 (0.57)
Personal significance	3.45 (0.88)	3.47 (0.75)	3.62 (0.80)	3.13 (0.85)	3.10 (0.93)	3.25 (0.98)	3.32 (0.72)	3.43 (0.83)
Valence	4.60 (0.46)	4.47 (0.51)	4.58 (0.47)	4.38 (0.60)	4.37 (0.57)	4.33 (0.44)	4.33 (0.50)	4.45 (0.58)
Active(1)/static event(2)	1.20 (0.25)	1.25 (0.31)	1.22 (0.31)	1.20 (0.28)	1.18 (0.31)	1.30 (0.39)	1.25 (0.34)	1.37 (0.37)
Self(1)/other perspective(2)	1.03 (0.13)	1.17 (0.36)	1.13 (0.29)	1.12 (0.25)	1.08 (0.19)	1.10 (0.24)	1.10 (0.20)	1.08 (0.27)
Recall frequency	3.23 (0.63)	2.78 (0.61)	2.95 (0.66)	2.55 (0.70)	2.72 (0.61)	2.88 (0.69)	2.63 (0.66)	2.83 (0.65)
Consistency	4.28 (0.49)	4.08 (0.76)	3.83 (0.75)	3.93 (0.54)	4.02 (0.65)	3.98 (0.53)	3.85 (0.71)	3.73 (0.69)
B. Objective scores								
Internal details	17.60 (5.42)	14.95 (3.95)	15.37 (5.96)	14.72 (6.75)	15.60 (4.84)	14.93 (5.74)	14.43 (3.89)	13.65 (4.50)
External details	6.35 (3.82)	6.03 (3.34)	6.38 (3.75)	6.08 (3.53)	5.82 (2.56)	6.02 (2.83)	5.78 (2.84)	6.07 (3.42)

Supplemental Table 2, related to Figures 6 and 7: Behavioural data (mean, SD) - Experiment 2 (n=16)

A. Subjective ratings	(0.5M) 8M	(4M) 12M	(8M) 16M	(12M) 20M	(16M) 24M	(20M) 28M	(24M) 32M	(60M) 68M
Vividness	3.47 (0.72)	3.41 (0.84)	3.34 (0.68)	3.19 (0.79)	3.09 (0.66)	3.31 (0.77)	3.22 (0.71)	3.09 (0.80)
Detail	3.22 (0.63)	3.25 (0.84)	3.13 (0.56)	2.84 (0.65)	3.06 (0.66)	3.09 (0.76)	2.97 (0.62)	2.75 (0.80)
Difficulty	2.34 (0.60)	2.19 (0.85)	2.22 (0.82)	2.28 (0.58)	2.25 (0.63)	2.22 (0.75)	2.28 (0.97)	2.25 (0.61)
Personal significance	2.97 (0.69)	3.13 (0.74)	3.22 (0.66)	2.78 (0.75)	2.84 (0.87)	2.94 (0.87)	2.66 (0.79)	3.00 (0.86)
Valence	4.31 (0.54)	4.28 (0.63)	4.28 (0.60)	4.06 (0.75)	4.22 (0.68)	4.03 (0.74)	3.97 (0.56)	4.38 (0.65)
Active(1)/static event(2)	1.31 (0.31)	1.41 (0.42)	1.38 (0.34)	1.47 (0.43)	1.19 (0.31)	1.34 (0.47)	1.34 (0.40)	1.28 (0.31)
Self(1)/other perspective(2)	1.09 (0.20)	1.13 (0.29)	1.13 (0.29)	1.16 (0.30)	1.00 (0.00)	1.06 (0.17)	1.16 (0.24)	1.09 (0.27)
Recall frequency	2.03 (0.69)	2.00 (0.80)	1.78 (0.75)	1.72 (0.45)	1.84 (0.77)	1.81 (0.70)	1.50 (0.48)	1.53 (0.59)
Consistency	3.94 (0.36)	3.94 (0.44)	3.91 (0.61)	3.88 (0.67)	3.91 (0.55)	4.00 (0.58)	3.81 (0.63)	3.72 (0.71)
B. Objective scores								
Internal details	18.75 (7.77)	17.56 (6.83)	17.59 (7.33)	18.66 (7.88)	19.63 (7.72)	20.31 (8.00)	17.06 (6.53)	16.03 (6.69)
External details	6.25 (4.27)	5.66 (4.06)	5.91 (2.78)	6.47 (4.46)	7.16 (3.34)	9.28 (5.40)	8.41 (5.18)	6.66 (4.36)