Classification: Biological Science – Neuroscience

Aging and the encoding of changes in events:

The role of neural activity pattern reinstatement

David Stawarczyk[1, 2], Christopher N. Wahlheim[4], Joset A. Etzel[1], Abraham Z. Snyder[3,5], Jeffrey M. Zacks[1,3]

[1] Department of Psychological & Brain Sciences, Washington University in St. Louis, St. Louis, MO 63105, USA.

[2] Department of Psychology, Psychology and Neuroscience of Cognition Research Unit, University of Liège, 4000 Liège, Belgium.

[3] Mallinckrodt Institute of Radiology, Washington University in St. Louis, St. Louis, MO 63110, USA.

[4] Department of Psychology, University of North Carolina at Greensboro, Greensboro, NC 27412, USA.

[5] Department of Neurology, Washington University in St. Louis, St. Louis, MO 63105, USA.

Correspondence concerning this article should be addressed to David Stawarczyk, Department of Psychology, University of Liège, Quartier Agora, place des Orateurs 1, 4000 Liège 1, Belgium. E-mail: d.stawarczyk@uliege.be. Tel.: 0032 4 3665979.

Abstract (250 words)

When encountering new events, memories of relevant past experiences can guide expectations about what will happen. When unexpected changes occur, this can lead to prediction errors, with consequences for comprehension and subsequent memory. For example, if a supermarket's produce section were moved after one's first visit, this could generate a prediction error, which could in turn drive better encoding of the new location. Aging could potentially impair the encoding of previous experiences, the retrieval of relevant instances, and their use in detecting and encoding changes. Using functional MRI multivariate pattern analysis, we investigated these mechanisms in healthy young and older adults. In the scanner, participants first watched a movie depicting a series of everyday activities in a day of the actor's life. They next watched a second movie of the day's events in which some scenes ended differently. Crucially, before watching the last part of each scene, the movie stopped, and participants were instructed to mentally replay the way the activity ended previously. Three days later, participants were asked to recall the activities. Individual differences in neural activity pattern reinstatement in the posteromedial cortex during the mental replay phases of the second movie were associated with better memory for changed features in the young but not older adults. This finding suggests that the posteromedial cortex contributes to the comprehension of perceived changes through reinstating previous event feature and that older adults are less able to use reinstatement to overtly recognize changes.

Keywords: Representational similarity analysis, cognitive aging, event cognition, episodic memory, change comprehension

Significance Statement (120 words max)

Change comprehension (the capacity to detect, encode, and later remember changes) is a critical ability supporting everyday functioning and, like many cognitive processes supported by long-term memory, could be particularly vulnerable to aging. Here, we examined the role of neural reinstatement of episodic memories in change comprehension using multivariate pattern-based functional MRI during viewing of movies of everyday activities. Detecting and remembering event changes were both associated with reinstatement in posteromedial brain areas of the episode-specific neural activity pattern present while viewing the original event. This effect was smaller in older adults. These results show how episodic retrieval can inform ongoing event comprehension and how this process is more effective for young than older adults.

Introduction

Why do humans and other animals remember? One important reason is that features of past experiences can guide current behavior. A key question is how humans use representations of past experiences to adapt to new situations. Recent proposals suggest that a critical function of event memory (1)—also referred to as episodic memory (2–4)—is to guide anticipation of upcoming events (5–9). One of these models, Event Memory Retrieval and Comparison Theory (EMRC; 10) proposes that associative retrieval of recent experiences drives predictions about how a current activity will unfold. In most cases, using event representations of past experiences facilitates predictions in similar new situations. However, when events unexpectedly change, memory-based predictions can lead to errors. Such errors impose a short-term cost but may have long-term benefits for detecting and registering that features of the environment have changed, as well as for encoding the new event features.

For example, suppose you visit a grocery store to pick up a gallon of milk, and then return a week later for more milk—only to find that the milk has been moved to a different refrigerated shelf. On your next visit, if you can retrieve the features of the second trip and their discrepancy from the first, you will be able to navigate efficiently to the new milk shelf. EMRC proposes that, during ongoing event comprehension, features of the present situation cue retrieval of similar event representations from long-term memory. These memory representations guide the formation of a working event model of the current situation, which generates predictions about what should happen next (11). When the features of the present event do not match these model-based predictions, a prediction error is experienced, leading to the updating of the current event model (12, 13). In such cases, the updated model can integrate features of the initial event, the mismatch between the event model-based predictions and ongoing perceptions, and the unexpected features of the current event. The formation of this configural memory representation can enable later retrieval of the original features, the changed features, and their temporal relations (14). Retrieval of such representations can allow more effective behavior in the future and constitute powerful opposition against memory interference. In the above-mentioned grocery story, it would for instance allow you to discriminate where the milk was located on the first visit from its location on the second visit. However, if during the second visit the encoding system does not register the discrepancy between the predicted and perceived features, it may form two separate memory representations without indication of their temporal relation, leading to interference (15).

The capacity to detect, encode, and later remember changes might be vulnerable to aging. There is ample evidence that episodic memory retrieval is impaired in healthy older adults, particularly when contextual information associated with previous experiences must be recollected (16, 17), and that this deficit extends to memory for naturalistic events (18, 19). In addition, prior studies using word lists have shown that older adults detect and later remember fewer changes than their younger counterparts (20), a finding corroborated by subjective memory complaints (e.g., difficulties in remembering the location of misplaced objects (21)) that fits well with the susceptibility to interference (22) and impairments in associative binding of event features that is characteristic of this age group (23, 24). From the perspective of EMRC, these age differences suggest that deficits in the ability to form configural memory representations by binding details of retrieved memories with perceived events might explain some of the difficulties with memory updating that older adults experience (25), and that this might impair the encoding of changes in events.

Evidence suggesting that older adults are less able to integrate details of retrieved memories during the comprehension of event changes comes from recent behavioral experiments using movie clips of everyday activities (10). In these experiments, viewers' ability to detect that a change had occurred and to remember this fact later was associated with better memory for the changed features. Older adults detected and remembered fewer event changes, and this was associated with greater memory disruption when a change occurred.

The current study tested a key proposal of EMRC: When perceiving a new activity, features of a related previous activity are retrieved and incorporated into the working event model, thus leading to predictions about how the new activity will unfold. In the current study, we aimed to directly assess the role of retrieving episode-specific event features when encoding a new event that was similar to an earlier event. To do so, we used functional MRI (fMRI) in combination with representational similarity analysis (RSA; 26) to assess whether the reinstatement of brain activity patterns associated with past events can facilitate change comprehension during the perception of new events.

Several neuroimaging studies have shown that patterns of brain activity present while encoding new information are reinstated when this information is recollected (e.g., 27–29). Although initially shown with classic laboratory stimuli such as word lists, similar evidence is obtained for more complex stimuli such as movies of everyday activities. For instance, recent studies have shown that reinstatement of brain activity patterns during recollection predicts retrieval accuracy of movie content up to one week after viewing (30–32). This effect is usually the strongest in the posterior areas of the default network (DN; 33, 34) and more specifically in the posteromedial cortex (PMC) that includes the posterior cingulate cortex (PPC) and retrosplenial cortex (Rsp), and in the medial temporal lobe (MTL) including the parahippocampal cortex (PHC) and hippocampus. These regions are sometimes referred to as the posterior medial system (35, 36) or contextual association network (37–39) due to their strong involvement in long-term memory recollection, particularly when episodic representations of everyday events must be remembered from visual cues (40).

The MTL and PMC might play key roles in encoding and remembering changes. Evidence suggests that the hippocampus and adjacent areas play a major role in encoding relations among perceptual features and comparing such relational information to information stored in memory in order to form configural memory representations of everyday experiences (41, 42). The hippocampus shows large metabolic alterations and volume loss in aging (43), making functional change in the MTL a potential source of age-related differences in episodic memory processes that enable change comprehension. The PMC, and in particular the PCC, has been associated with supporting event model representations (9, 44, 45). Like the MTL, the PMC undergoes substantial metabolic and structural change in aging (46), with the integrity of its functioning being related to better cognitive abilities in older adults (47). These considerations make the MTL and the PMC strong candidates for encoding relational information during event comprehension and supporting the reinstatement of such information when current events are similar to previously encountered situations. However, there is no evidence to date establishing the reinstatement of brain activity patterns during the comprehension of changes in events, nor is there evidence regarding how pattern reinstatement in these regions differ between older and young adults.

To address these issues, we used a task similar to the one used in (10). Healthy young and older adult participants viewed two movies depicting everyday activities in two days of an actor’s life (hereafter referred to as Day 1 and Day 2) during fMRI. The Day 2 movie depicted activities that were either repeated exactly or began the same but ended differently (**see Figure 1**). In the present study, we stopped each activity in the Day 2 movie a few seconds after its onset (i.e., before any change had begun) and asked participants to mentally replay the activity ending of the Day 1 movie. After this reinstatement phase, participants then saw the ending of the activities where the changes occurred. They were asked after the ending if they could remember what happened in the Day 1 movie when the Day 2 movie stopped, and whether the end of the clip that followed included a repeated or changed feature. We then used fMRI multivariate pattern analysis to determine the extent to which brain activity patterns in PMC and MTL were similar between the Day 1 viewing of activity endings and Day 2 reinstatement phase by computing a reinstatement score for each activity and participant. Three days later, participants were given an unscanned cued recall test for the activities of the Day 2 viewing.

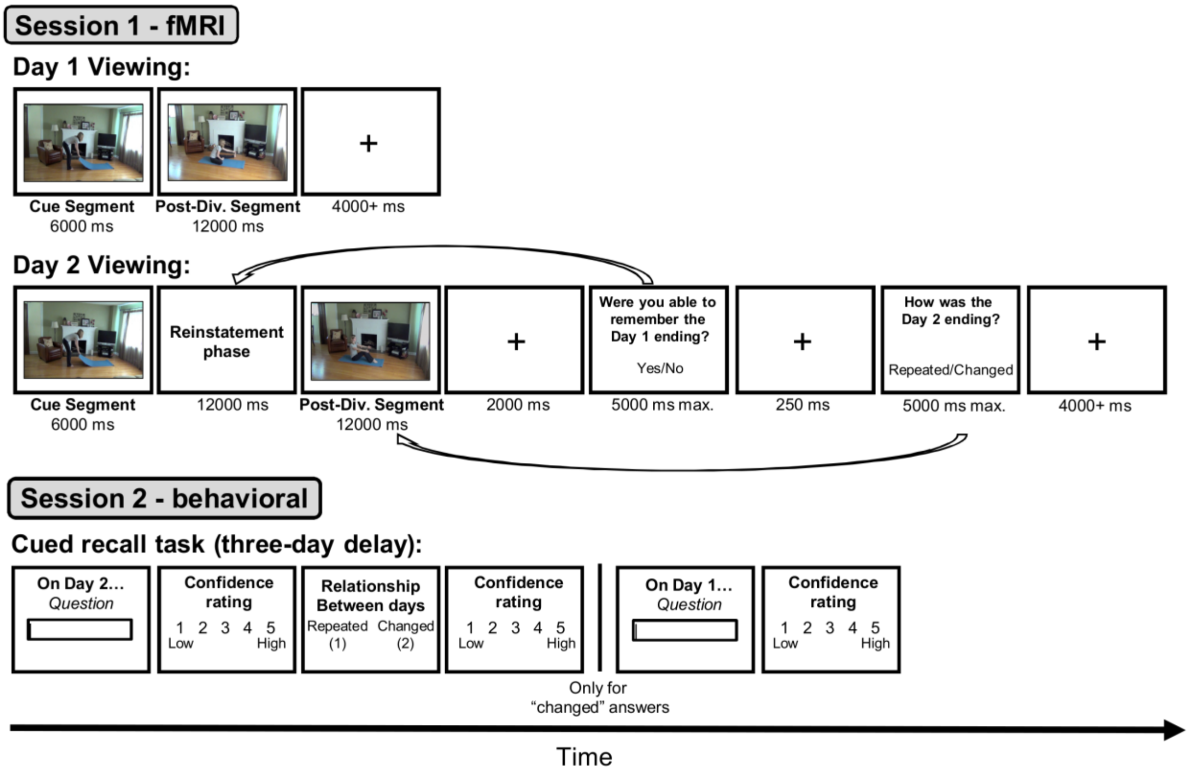


Figure 1. Trial structure of the tasks. During the Day 1 viewing, the transition from the cue to the post-divergence segment of each activity appeared seamlessly. During the Day 2 viewing, the activities were interrupted at the end of the cue segment by a reinstatement phase where participants were asked to mentally replay the corresponding post-divergence segment from the Day 1 viewing. All the cue segments during Day 2 were identical to those of Day 1. Fifteen of the post-divergence segments were repeated exactly from Day 1 while the remaining 30 included a changed feature (as in the illustrated example). The cued recall task took place outside the scanner three days after the first session. The questions about Day 2 and Day 1 were specific to each activity and were always focused on the critical feature that changed between the two versions of each activity (e.g., "What did the actor do on the exercise mat?"). The duration of every step in this task was self-paced. For a more detailed description of the materials and procedure, see the Supporting Information.

We hypothesized that stronger reinstatement of MTL and PMC activity patterns present while watching the Day 1 movie would be associated with better detection of change during the Day 2 viewing, better subsequent recall of the changed features, and better subsequent recollection that a change had occurred (including what the changed feature had been during the first viewing). Furthermore, given the decline of MTL and PMC functional integrity with aging, we expected that these associations between reinstatement and change detection, memory for changed activity features and memory for change itself would be stronger for young than older adults and that this difference might be explained (at least partly) by older adults' diminished ability to retrieve event memories and create configural memory representations.

Results

All analyses of memory performance and reinstatement effects were computed using mixed effect models with subjects and activities as random effects. Logistic models were used when the dependent variable was binary. Mixed effects models were fitted using the lme4 package in R (48), hypothesis tests were performed using the Anova function of the car package (49), and post hoc comparisons using the Tukey method were conducted using the emmeans package (50). Comparisons between nested models were performed with the anova function of the lme4 package. Finally, the plot_model and get_model_data functions from the sjPlot package (51) were used to estimate the model predicted values for description and visualization.

Behavioral results

We first examined recall of repeated and changed Day 2 activity features (**Figure 2**). A 2 (Activity Type: Repeated vs. Changed) × 2 (Age: Young vs. Older) model revealed that changed activities were recalled less accurately than repeated activities [$\chi 2(1) = 25.91$, $p < .001$]. Older adults recalled fewer Day 2 activities than young adults [$\chi 2(1) = 9.61$, $p = .002$]. There was no significant Activity Type × Age interaction [$\chi 2(1) = 0.37$, $p = .54$]. Regarding the ability to classify changed activities as such (Table 1, top rows), a model with Age as a fixed effect revealed no significant difference between the two groups [$\chi 2(1) = 1.80$, $p = .18$].

However, recall of Day 1 activity features following correct change classifications (Table 1, bottom rows) was greater for young than older adults [$\chi 2(1) = 5.11$, $p = .02$].

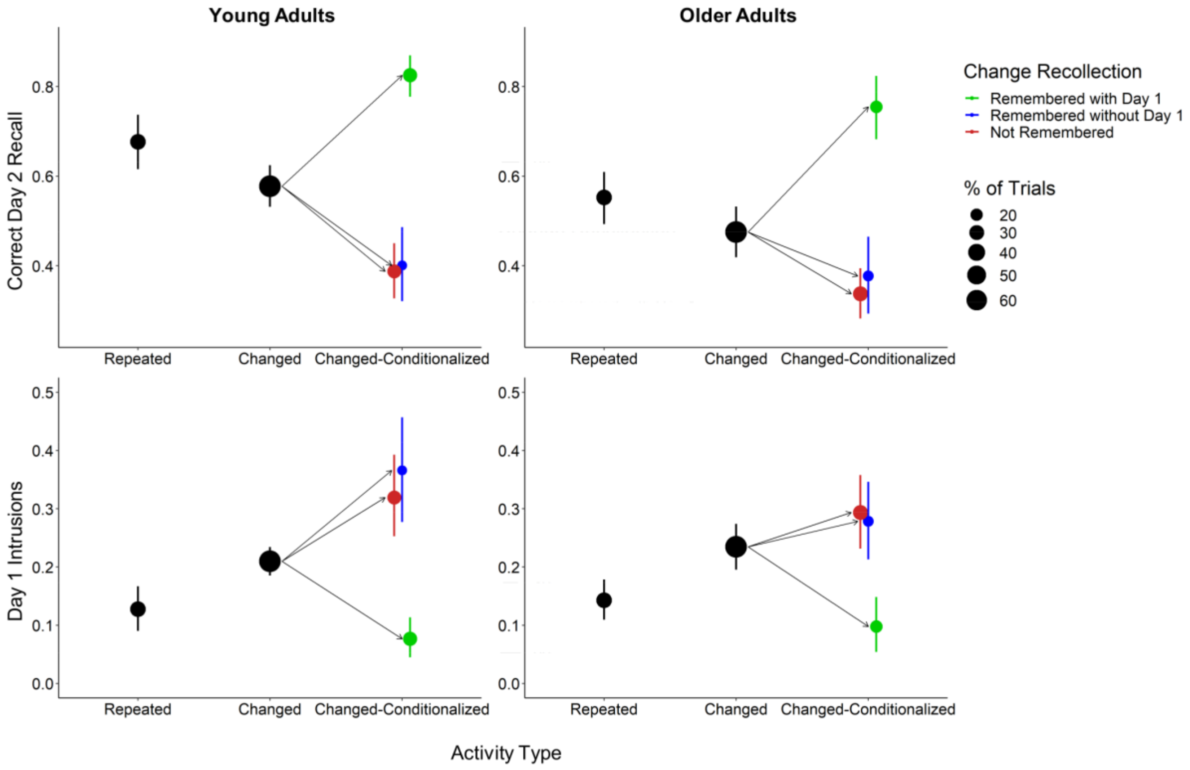


Figure 2. Mean probabilities of correct Day 2 recall and Day 1 intrusions computed over all the trials of each group. Error bars are bootstrap 95% confidence intervals. (Note that intrusions for Repeated activities are cases where participants guessed an ending that they never saw.)

Table 1: Model predicted values for memory for change and Day 1 recall accuracy for changed activities correctly remembered as changed based on age group.

	Young adults	**Older Adults**
Memory for change	.63 [.54, .70]	.55 [.46, .64]
Day 1 Recall	.69 [.60, .77]	.55 [.44, .65]

Note: 95% confidence intervals are displayed in brackets.

We then examined differences in Day 2 recall for changed activities associated with remembering change and recalling Day 1 features compared with Day 2 recall for repeated activities (**Figure 2, top panels**). The model for Day 2 recall accuracy of the activity features including repeated and changed activities conditionalized on change classifications and Day 1 recall for both age groups indicated an effect of Activity Type [$\chi 2(3) = 229.03$, $p < .001$] and no significant interaction with Age [$\chi 2(1) = 2.13$, $p = .54$]. Post-hoc tests showed that recall of changed features when participants remembered that an activity had changed and correctly recalled the Day 1 feature was higher than recall of repeated features (*z* ratio = 5.95, $p < .001$). Day 2 recall accuracy for changed activity features when change was not remembered or remembered without the Day 1 feature did not differ from each other (*z* ratio = 0.29, $p = .99$) and were both lower than recall accuracy for repeated activity features (smallest *z* ratio = -7.64, $p < .001$).

Analyses of Day 1 intrusions during Day 2 recall (**Figure 2, bottom panels**) revealed results that mirrored those of correct recall. The model including repeated activities and conditionalized change activities for both age groups indicated an effect of Activity Type [$\chi 2(3) = 144.00$, $p < .001$] and no interaction [$\chi 2(3) = 2.26$, $p = .52$]. For both age groups, remembering that an activity had changed and recalling the Day 1 feature was associated with lower intrusions compared to baseline intrusion rates for repeated items (*z* ratio = -3.10, $p = .01$). This is not surprising given that the observations for changed activities were instances when participants reported the Day 1 feature twice, presumably because they were guessing. Intrusion rates for changed activities when changes were not remembered or when they were remembered but the Day 1 feature was not recalled were not significantly different (*z* ratio = -0.92, $p = .79$). Both conditional cells were associated intrusion rates above repeated activities (smallest *z* ratio = 7.56, $p < .001$). Given the lack of difference in Day 2 recall performance when change was not remembered or remembered without the Day 1 feature, we collapsed across these cells in subsequent analyses. Conditional cells for change activities were classified as *change recollected* (remembered as changed and the Day 1 feature correctly recalled) and *change not recollected* (not remembered as changed or remembered without the Day 1 feature).

In summary, results showed thus far that recollecting change was associated with better memory for changed activity features and that the magnitude of this effect was comparable for both age groups. However, older adults recollected the original feature of changed activities less often than young adults.

We next examined participants' self-reported success in reinstating Day 1 activity features during Day 2 viewing and their classification of activity types as repeated or changed while they viewed the Day 2 movie (**see Table 2**). A 2 (Activity Type: Repeated vs. Changed) × 2 (Age: Young vs. Older) model for reinstatement success indicated that older adults reported greater success in remembering Day 1 activity features than younger adults [$\chi2(1) = 4.55$, $p = .03$]. In contrast, a comparable model for activity classifications showed the opposite pattern: younger adults classified both changes and repetitions in the scanner more accurately than older adults [$\chi2(1) = 6.68$, $p = .01$]. No other effects were significant [*largest* $\chi2(1) = .30$, $p = .58$]. We examined the relationship between reinstatement success and change detection using a 2 (Reinstatement: Successful vs. Unsuccessful) × 2 (Age: Young vs. Older) model (**see Table 3**). The model indicated a significant Reinstatement × Age interaction [$\chi2(1) = 9.97$, $p = .002$], showing that reinstatements reported as successful were associated with more accurate change detection for young adults (z ratio = 5.94, $p < .001$) but not for older adults (z ratio = 1.00, $p = .32$).

Table 2: Model predicted values for self-reported reinstatement and activity classification accuracy from the Day 2 viewing based on activity type and age group.

		Activity Type	
Dependent variable	**Age group**	**Repeated**	**Changed**
Self-reported reinstatement	Young	.80 [.71, .87]	.79 [.70, .86]
	Older	.89 [.83, .94]	.88 [.81, .93]
Activity Classifications	Young	.87 [.81, .91]	.88 [.83, .91]
	Older	.79 [.71, .85]	.80 [.73, .85]

Note: 95% confidence intervals are displayed in brackets.

Table 3: Model predicted values for change detection accuracy based on self-reported reinstatement accuracy and age group.

	Self-reported reinstatement	
Age group	**Successful**	**Unsuccessful**
Young	.93 [.89, .95]	.79 [.69, .86]
Older	.82 [.75, .88]	.79 [.67, .86]

Note: 95% confidence intervals are displayed in brackets.

In addition, we also examined whether self-reported reinstatement accuracy and change detection while viewing the Day 2 movie in the scanner were related to performance in the cued recall task, and whether self-reported reinstatement and change recollection remained significant predictors of Day 2 recall accuracy for the changed features above and beyond each other. These analyses are reported in the Supporting Information. In summary, they showed that self-reported reinstatements of Day 1 activity features during Day 2 viewing were associated with better change comprehension. However, these benefits were smaller for older adults: Self-reported reinstatement was not associated with reduced Day 1 intrusions or increased change detection in this age group. Finally, the more accurate recollection of changed features observed when participants reported successfully reinstating the Day 1 activity features did not remain significant after controlling for the effect of change recollection.

RSA results

To assay the neural reinstatement of activity-specific Day 1 features during Day 2 viewing, we estimated the similarity of the activity pattern during the viewing of each activity's ending on Day 1 to that during the attempted recollection of each ending on Day 2 during the reinstatement phases (see Figure 1). Similarity was computed between all possible pairings of Day 1 and Day 2 activity patterns using Pearson correlation across the voxels within each region of interest. Finally, we computed a reinstatement score for each activity, which describes the tendency for Day 1 and Day 2 patterns from viewing/recalling the same activity to be more positively correlated than those from viewing/recalling different activities (32, 52). Specifically, the reinstatement scores are the correlations for same-activity pairings minus the mean correlation of all the non-matching pairs for that participants.

We first computed mean reinstatement scores for the MTL and PMC by averaging for each changed trial the reinstatements scores of all the parcels within each of these two areas from the DN subsystems of the 17 networks/300 parcels cortex parcellation map built by Schaefer et al. (53) (see **Figure S1**). To determine whether there were age differences in reinstatement scores, we fitted models with mean reinstatement score as the dependent variable, participants and activities as random effects, and age group as fixed effect. The analyses showed no significant differences between the two age groups in either the PMC or MTL [$\chi^2(1) = 0.55$, $p = .56$, for the PMC; $\chi^2(1) = 1.29$, $p = .26$, for the MTL] (see **Figure 3**). The same models fitted on the reinstatement scores for each individual parcel were not significant except for parcel 144 in the left parahippocampal cortex for which reinstatement scores were higher for the young than older adults (see **Table S1**). For descriptive purposes, we also calculated overall reinstatement scores for all parcels in the atlas (**Figure S2**). This revealed evidence of widespread reactivation throughout the cortex. All subsequent analyses were restricted to the *a priori* ROIs.

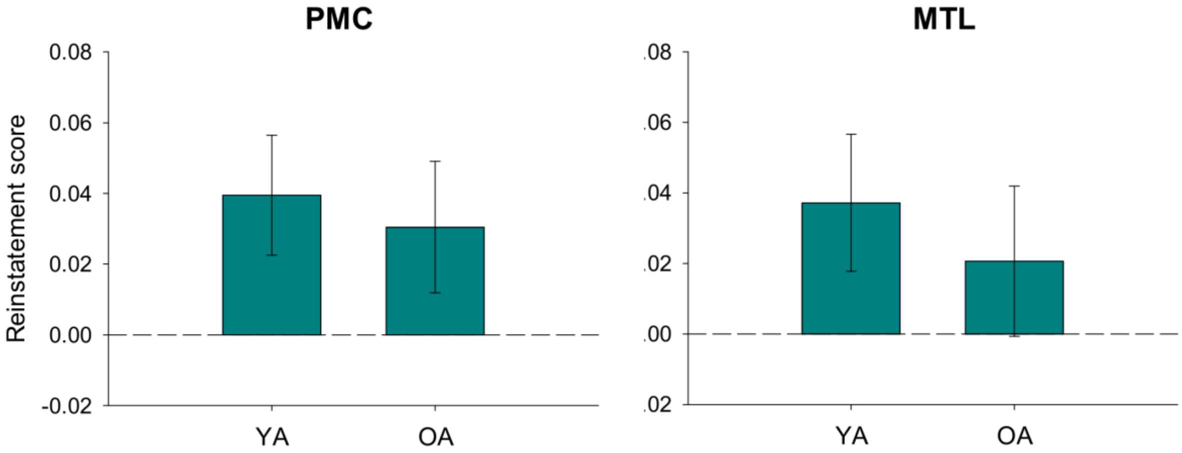


Figure 3. Predicted values for mean reinstatement scores in the MTL and PMC by age group. The error bars represent 95% confidence intervals. YA = Young adults; OA = Older adults.

Next, to examine within-individual relationships between reinstatement and change comprehension measures, we group mean-centered the reinstatement scores for the changed activities of each participant, which allowed us to examine within-individual differences. To examine between-individual differences, we created a new variable that corresponded to the mean reinstatement score for the changed activities of each participant replicated across all changed trials for that participant that we grand-mean centered (54). We then fitted models with the between-individual and within-individual reinstatement scores, as well as their respective interaction with Age as a fixed effect. In addition, because reinstatement scores across parcels were only moderately correlated (**Table S2**), we performed similar analyses but examined the effect of each individual PMC and MTL parcel on memory performance above and beyond the effects of all the other parcels in the area (see the Supporting Information for details).

At the between-individual level, participants with higher mean PMC reinstatement scores had higher Day 2 recall performance for changed activities [$\chi 2(1) = 4.72$ $p = .03$] and this effect did not interact with Age [$\chi 2(1) = 2.01$ $p = .16$]. For the MTL, the effect of mean reinstatement score at the between-individual level effect did not reach statistical significance [$\chi 2(1) = 3.54$ $p = .06$] but the interaction with Age did [$\chi 2(1) = 4.09$ $p = .04$], indicating that the association between reinstatement scores and recall of the changed features was weaker in the older compared to the young adults. Models computed separately for each age group showed that higher mean MTL reinstatement scores was associated with better recall in the young [$\chi 2(1) = 8.61$ $p = .003$] but not older adults [$\chi 2(1) = 0.18$ $p = .67$]. No significant effect was found at the within-individual level, and analyses of Day 1 intrusions revealed no significant effects of either the mean PMC or MTL reinstatement scores and no significant interaction with Age at either the between- or within-individual level (all *ps* > .05). The model predicted values from the between-individual reinstatement terms for each dependent variable are presented in **Figure 4**; the corresponding values predicted from the within-individual reinstatement terms are presented in **Figure S3**.

We next examined whether RSA reinstatement scores were related to change recollection accuracy. As for correct Day 2 recall, the mean PMC reinstatement score was again significant at the between-individual level [$\chi 2(1) = 4.17$ $p = .04$], but this time the effect was qualified by an interaction with Age [$\chi 2(1) = 5.26$ $p = .02$], indicating that the association between reinstatement scores and change recollection was weaker in the older than young adults. Follow-up analyses showed that higher mean reinstatement scores were related to higher change recollection accuracy in young [$\chi 2(1) = 8.36$ $p = .003$] but not older adults [$\chi 2(1) = 0.31$ $p = .86$]. Similar results at the between-individual level were found for the MTL with a significant effect of mean reinstatement score [$\chi 2(1) = 4.98$ $p = .03$] and significant interaction with Age [$\chi 2(1) = 3.86$ $p = .049$]. Young people with higher reinstatement scores recollected more changes [$\chi 2(1) = 7.24$ $p = .007$], but this relationship was not significant for older adults [$\chi 2(1) = 0.54$ $p = .46$] (see **Figure 4**). Again, no significant effect was found at the within-individual level for either the mean PMC or MTL reinstatement scores (all *ps* > .05; see **Figure S3**).

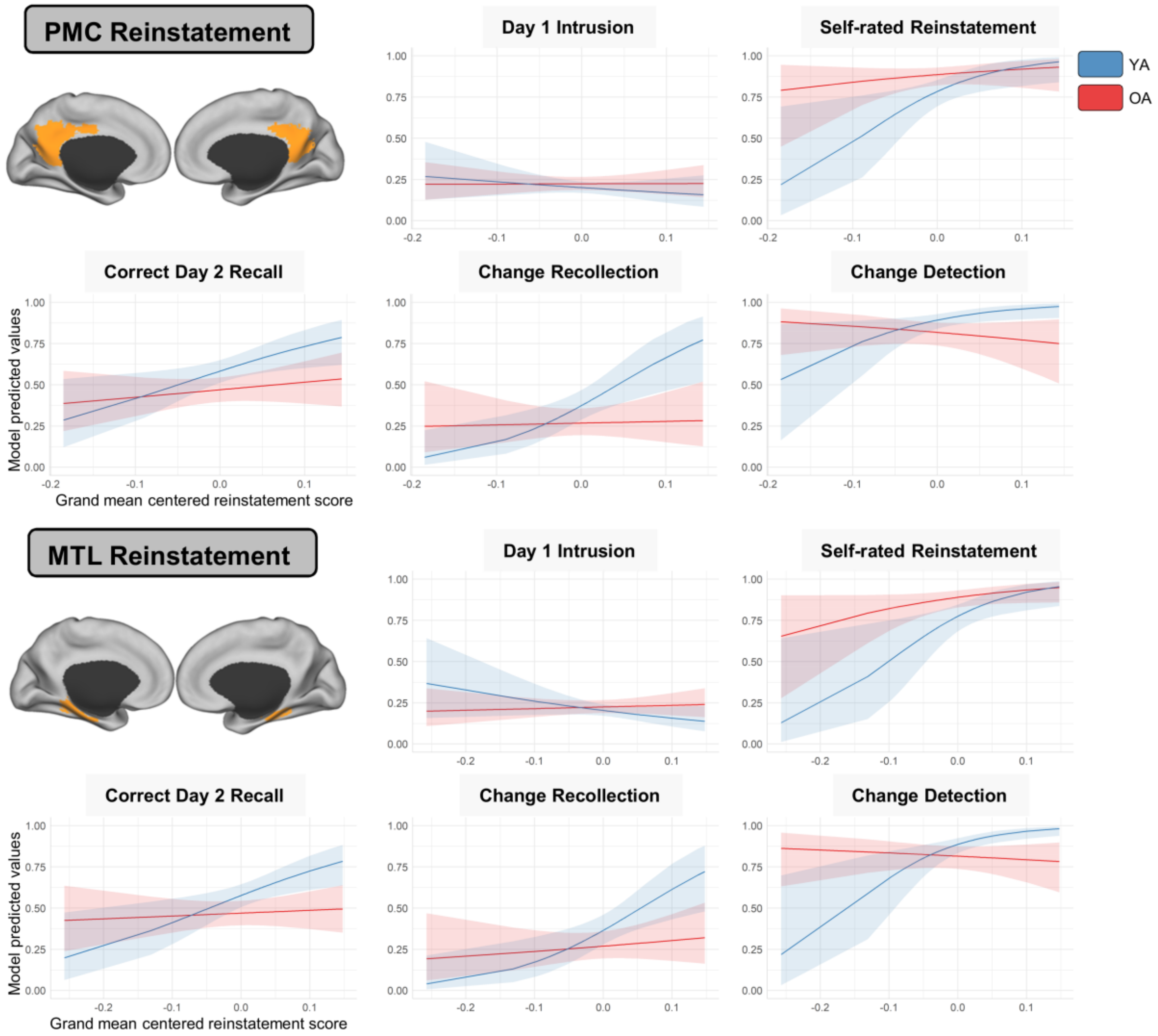


Figure 4. Predicted values and 95% confidence intervals for the association between change comprehension measures and mean PMC/MTL reinstatement scores (between-individual level). YA = Young adults; OA = Older adults.

We then examined whether RSA reinstatement scores were predictive of self-reported reinstatement and change detection accuracy in the scanner. Regarding self-reported reinstatement success, participants with higher mean PMC reinstatement scores rated more of their reinstatements as correct [$\chi^2(1) = 5.00$, $p = .03$], and this effect did not interact with Age [$\chi^2(1) = 2.11$, $p = .15$]. A similar effect of mean reinstatement score at the between-individual level was found for the MTL [$\chi^2(1) = 8.40$, $p = .004$] and did not significantly interact with Age either [$\chi^2(1) = 1.32$, $p = .25$] (see **Figure 4**). Next, for change detection accuracy, the effect of mean reinstatement scores in the PMC and MTL were both qualified by significant interactions with Age [$\chi^2(1) = 5.57$, $p = .02$, for the PMC; $\chi^2(1) = 8.98$, $p = .03$, for the MTL], indicating that the association between reinstatement scores and self-reported reinstatement success was weaker in the older than young adults. Follow-up analyses for both the PMC and MTL showed that higher mean reinstatement scores were predictive of more accurate change detection for young [$\chi^2(1) = 4.76$, $p = .03$, for the PMC; $\chi^2(1) = 9.81$, $p = .003$, for the MTL], but not older adults [$\chi^2(1) = 0.81$, $p = .36$ for the PMC; $\chi^2(1) = 0.27$, $p = .60$ for the MTL]. No significant effect of mean reinstatement scores or interaction with Age was found at the within-individual level for either self-reported reinstatement or change detection accuracy (all *ps* > .05, see **Figure S3**).

Next, because between-individual differences in mean PMC and MTL reinstatement scores were positively related to Day 2 recall and change recollection for young adults, we examined whether the association between reinstatement scores and Day 2 recall accuracy for the changed activities could be accounted for by change recollection. EMRC predicts that remembering previous event features while viewing new activities should enable the formation of configural memory representations and thereby facilitate memory for changes on a later test. These models thus included both Change Recollection (Recollected vs Not recollected) and mean reinstatement scores at the between-individual level as fixed effects (see Figure S4 for the predicted Day 2 recall accuracy values by mean reinstatement scores and Change Recollection above and beyond the effect of each other). Results revealed that neither mean PMC nor mean MTL remained significant terms above and beyond Change Recollection [$\chi 2(1) = 1.33$, $p = .25$, for the PMC; $\chi 2(1) = 3.25$ $p = .07$, for the MTL], whereas change recollection was still associated with higher Day 2 recall accuracy [$\chi 2(1) = 147.90$, $p < .001$, for the PMC model; $\chi 2(1) = 147.65$ $p < .001$, for the MTL model]; neither interaction term was significant (both *ps* > .05).

Finally, as both self-reported reinstatement and RSA reinstatement were associated with higher recall accuracy for the changed Day 2 features for young adults, and because neither of these effects remained significant after controlling for the effect of change recollection accuracy, we examined the contribution of self-reported reinstatement and RSA reinstatement on Day 2 recall accuracy for changed activities above and beyond each other (see the Supporting Information for details on these analyses). Results showed that both measures independently predicted Day 2 recall accuracy.

Discussion

The first goal of the present study was to test the proposal that reinstating features of a related previous activity can facilitate encoding of a new activity that includes a change by enabling the formation of a configural representation (10). The second goal was to investigate how these processes differ between healthy young and older adults. Consistent with our hypotheses, young adults who better reinstated neural activity patterns in the MTL and PMC just before encoding a changed activity ending remembered that activity better three days later. This effect was attenuated in older adults.

The positive association between neural pattern reinstatement and memory for changed activity features was statistically explained by the ability to recollect the fact that the activity had changed and the original event feature. A similar pattern was seen for participants' introspective judgments about whether they had successfully reinstated event features before viewing each event's ending. However, self-reported and neural pattern reinstatement both independently predicted memory accuracy for the changed features, suggesting that these two measures reflect partly distinct processes.

Neural pattern reinstatement in the PMC and MTL was associated with a more accurate recollection of new information presented *after* the end of the reinstatement phase that *conflicted* with the previously encoded features that were reinstated. This finding is consistent with previous studies showing that reinstatement, during recollection, of the brain activity pattern present in posterior DN areas while watching movies can predict memory for the movie content up to one week later (30–32). The current results indicate for the first time that such reinstatement is related to the encoding of novel, unexpected event features. According to EMRC, this effect results from prediction errors that are triggered by the mismatch between event model-based predictions generated from memory representations of the original event features and perceptual inputs gathered when experiencing the changed event (10). One proposed benefit of experiencing such errors is the creation of configural representations composed of the original activity features, the changed features, and their temporal relations. The present results support this view because neither self-reported nor neural pattern reinstatement remained significant predictors of memory accuracy for the changed features after controlling for change recollection accuracy. This suggests that the beneficial effect of pattern reinstatement on remembering changed activities features is mediated by the retrieval of configural memory representations.

The PMC may play a key a role in supporting the event model representations from which predictions are generated. The PMC is part of the DN, and it was initially thought to be exclusively involved in generating internal mentation that stands in opposition with attention to the external world (55). However, there is now substantial evidence that the PMC also supports externally-directed attention and event comprehension when task performance and the processing of perceptual inputs can benefit from relevant information stored in memory (52, 56). In addition, recent studies have revealed that the transition between activity patterns within the PMC while watching movies follow a timescale ranging from seconds to minutes that closely matches how people segment movie content into distinct events (44), supporting the view that event model representations might be the means by which PMC facilitates interactions with the external world. As for the MTL, there is extensive evidence that this region is involved in the relational binding of information stored in memory and how it relates to perceptual inputs in order to form associative memory representations of everyday experiences (41, 42). Consistent with these findings, recent research suggests that peaks of activity in the hippocampus at the transition between perceived events can predict neural reinstatement in the PMC during recall (44, 57). Pattern reinstatement within the MTL in the current study might thus reflect the relational binding of information stored in memory—whose retrieval is triggered by the cue segment—in order to form the event model that is supported by the PMC. This event model would then help to detect and encode the changes that are experienced when ongoing experiences do not match model-based predictions.

Next, although the overall strengths of neural activity pattern reinstatement scores in the PMC and MTL did not differ between young and older adults, their relationship to memory did. The relationship between neural reinstatement at the between-individual level and change comprehension was weaker—indeed, largely absent—in older adults. In addition, older adults were less likely to detect changes and recall the original activity features during the cued recall task than their younger counterparts, suggesting an age-related deficit in the ability to establish configural representations. This fits well with previous findings on change comprehension (10, 20) and with well-established age-related deficits in associative memory (23, 24). A possible explanation for these findings is suggested by behavioral studies showing that event perception is impaired in aging (18, 19): The event representations formed by the older adults during the original viewing of the activities—and that were retrieved during the reinstatement phases—may have been less detailed than those of younger adults, resulting in event models that were possibly coarser, less coherent, and thus less likely to facilitate change comprehension. Although hypothetical, this proposal might explain why self-reported reinstatement accuracy in older adults did not predict better change detection and was not associated fewer Day 1 intrusions nor correct recognition of the changed activity features, as was the case for the young adults. Further studies examining the features of memories retrieved during the reinstatement phase (for instance by asking participants to verbalize the retrieved activity content) would be helpful to determine how aging affects the informational content of event features retrieved during ongoing comprehension. In any functional neuroimaging study comparing young and older adults, it is important to consider potential sources of artifact; these include group differences in neurovascular coupling, in head motion, and in how the tasks are approached (58). In the present case, the fact that older adults showed robust overall neural reinstatement renders their significantly weaker relationships between neural reinstatement and behavioral memory measures particularly significant.

In conclusion, the present results showed that the reinstatement of previously-seen related events can facilitate the encoding of features that one is about to encounter that include changes from prior events. This is particularly striking because the new features conflict with the just-retrieved features of the previous activity. We proposed that retrieving activity features facilitates this encoding precisely because it enabled registering a discrepancy between the predicted event features and the encountered features. This process was impaired in older adults; the pattern of impairment suggested that deficits in encoding a detailed memory representation of the original event might reduce older adults' ability to encode a configural representation of the changed event that includes its relationship to the previous event. More generally, these results support the recent view that DN activity does not exclusively support the generation of internal mentation that competes for cognitive resources with processes that enable attention to the external world (55, 59), but can also facilitate the processing of upcoming events, possibly through predictive processes based on events models that integrate past event representations with current perceptual inputs (52).

Method

The full stimulus sets for the materials used in the present experiments, anonymized data files, coded data, and R Markdown files (60) containing the analysis scripts are available on the Open Science Framework: (https://osf.io/v3dqg/). The research reported here was approved by the Institutional Review Board of Washington University in St. Louis.

Participants

The sample included 62 healthy right-handed participants: 34 young adults (mean age 22.85 years, *SD* = 2.71, range: 18-27 years, 22 females) and 28 older adults (mean age 69.86 years, *SD* = 5.01, range: 65-84 years, 20 females). All older adults had a score of 25 or above (*M* = 29.25, *SD* = 0.87, range: 27-30) on the Mini-Mental State Exam (MMSE; 61). For more details about recruitment and exclusion criteria, see the Supporting Information.

Materials, Design, and Procedure

The materials were movies of a female actor performing daily activities on two fictive days in her life (10). Two versions (A and B) of each activity that differed on a thematically central feature were used (e.g., doing stretching or sit-ups on a yoga mat, see **Figure 1**). The complete stimulus set consisted of 45 pairs of activities. The sequence of activities during each fictive day was fixed, similar between the two movies, and included of the the 45 activities—beginning with the actor waking up and ending with her going to bed. The version of the activity that participants saw in the first movie, and the 30 activities that were changed from the first to the second movie were counterbalanced across participants.

Participants completed the task in two sessions separated by three days (see **Figure 1**). During the first session, participants watched both movies while lying in the scanner. After participants viewed the Day 1 movie, field map images and a high-resolution anatomical image were collected, taking approximately 10 minutes. Participant then saw the second movie, we called this second functional run the "Day 2" viewing. At the end of this second functional run, we collected a second set of field map images and a high-resolution T2-weighted image, taking approximately 6 minutes. We then informed participants that that the session was over.

During Session 2, which took place outside of the scanner, we first tested participants' memory for the previously-viewed activities using a cued recall task (see **Figure 1**). During the task, the questions appeared in the same order as the activities during each movie. Following the cued recall task, we administered a recognition test of Session 1 activities (see the Supporting Information). After the recognition test, all participants completed a vocabulary test (62), and older adults completed the MMSE (61) last. All stimulus materials for the movies, cued recall task, and recognition task were presented using E-Prime 2 software (Psychology Software Tools, Pittsburgh, PA, USA; 3). A more detailed description of the procedure, materials, and instructions given to the participants is provided in the Supporting Information.

fMRI Data Analyses

Because we had strong *a priori* hypotheses regarding the brain regions that would be relevant in our RSA, we chose to use a ROI-based analytic strategy. Specifically, we selected the PMC and MTL parcels of the DN subsystems from the 17 networks/300 parcels cortex parcellation map of Schaefer et al. (53). We resampled the parcellation to atlas space using the 3dresample AFNI command and warped it from MNI to our template using the "Old Normalise" command in SPM 12. The resulting ROIs are presented in **Figure S2**. In addition, as the parcellation of Schaefer et al. (53) does not map subcortical areas, we added two parcels for the left and right hippocampus using the Oro.nifti package (64) in R. These hippocampal parcels were obtained from the minimal preprocessing pipelines for the Human Connectome Project (65) using the “Cifti Separate” workbench command.

Following spatial preprocessing (see the Supporting Information for details), data were normalized and detrended using second-order polynomials with the 3dDetrend AFNI command, spatially smoothed (66) with a Gaussian kernel of 3-mm full-width at half maximum in SPM, and z-scored with PyMVPA (67). To summarize the activity within each voxel during the period of interests in each run, we performed temporal compression by averaging the ninth to fourteenth scans (11.97 to 18.62 s) after the beginning of each activity using the 3dTstat AFNI command. This averaging window corresponded to the post-divergence segments in the first run and reinstatement phases in the second run. It was chosen to account for the hemodynamic delay, and to reduce the chance of contamination from the blood-oxygen-level-dependent (BOLD) response associated with the fixation cross of the first run and post-divergence segment of the second run, since these started 18 s after the onset of the cue segment. This temporal compression procedure resulted in one brain image for each activity, run, and participant.

We then compared the similarity of the brain activity patterns in each parcel between the two runs. Specifically, a similarity matrix was constructed for each parcel within each participant by placing the 45 activities from the first run along one axis and the same 45 activities from the second run along the other axis, then computing all possible pairwise Pearson correlations (i.e., each Day 1 post-divergence segment correlated with all Day 2 reinstatement phases). If participants successfully reinstated the pattern of brain activity they had while watching the post-divergence segments of the first run during the reinstatement phases of the second run, we would expect the correlations along the diagonal (i.e., watching and reinstating matching activity pairs) to be higher than the off-diagonal correlations (i.e., watching and reinstating different activity pairs). This tendency was quantified by computing a “reinstatement score” for each activity: the average of the off-diagonal cells subtracted from each of the cell along the diagonal (after Fisher’s r to z transformation; 31, 56), which reflects how specifically participants reinstated the pattern of brain activity they had while watching the post-divergence segments of the first run during the reinstatement phases of the second run.

Response coding

All open-ended responses made during the cued recall task (i.e, answers to the questions about the Day 2 activity features and questions about the Day 1 activity feature for activities classified as changed) were rated by the first author; a second independent judge also classified the responses to compute the interrater agreement. Participants’ responses during the cued recall test of Day 2 activities were assigned to one of three categories. *Correct Day 2 recall* responses were correct descriptions of the criterial activity feature in the Day 2 activity. *Day 1 intrusions* were responses that included the criterial feature from the version of the activity that was not shown on Day 2. All remaining responses were considered as *Incorrect*. The Cohen’s κ (68) between the two independent raters was .85. When participants reported on the cued recall test that an activity presented in the Day 2 viewing had changed from the Day 1 viewing, they were asked to recall what had happened on Day 1. These responses were coded as *correct Day 1 recalls* and *incorrect Day 1 recalls* by the same two independent raters depending on whether the participants recalled the critical activity feature in the Day 1 activity. Cohen’s κ for these ratings was .86.

Acknowledgements

We thank Ryan Kahle, Madeleine Schroedel, and Priscilla Mei for assistance with data collection and coding. We also thank Aaron B. Tanenbaum for his help with spatial preprocessing and selection of the fMRI sequences. This project was funded by NIH grant R21AG05231401 and supported in part by the Neuroimaging Informatics and Analysis Center (1P30NS098577). David Stawarczyk is currently supported by a Marie Curie Global Fellowship grant from the European Commission (Grant Agreement number: 798109). Joset A. Etzel was partially supported by the National Institutes of Health, grant number R37MH066078.

Author contribution

DS, JE, CW, and JZ designed the study. DS performed the study. DS, JE, and JZ analyzed the data. DS wrote the paper with all co-authors providing critical inputs. AS helped with the selection of the fMRI sequence and provided the spatial preprocessing pipeline used in the present work. All authors approved the final version of the manuscript.

References

1. D. C. Rubin, S. Umanath, Event memory: A theory of memory for laboratory, autobiographical, and fictional events. *Psychol Rev* **122**, 1–23 (2015).

2. E. Tulving, Episodic memory: from mind to brain. *Annu Rev Psychol* **53**, 1–25 (2002).

3. A. Baddeley, J. P. Aggleton, M. A. Conway, *Episodic memory: New directions in research* (Oxford University Press; US, 2002) https:/doi.org/10.1093/acprof:oso/9780198508809.001.0001.

4. E. Dere, A. Easton, L. Nadel, J. P. Huston, Eds., *Handbook of Episodic Memory* (Elsevier, 2008).

5. M. Bar, E. Aminoff, M. F. Mason, M. Fenske, The units of thought. *Hippocampus* **17**, 420–428 (2007).

6. L. W. Barsalou, Simulation, situated conceptualization, and prediction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **364**, 1281–1289 (2009).

7. D. L. Schacter, *et al.*, The future of memory: remembering, imagining, and the brain. *Neuron* **76**, 677–94 (2012).

8. S. B. Klein, The temporal orientation of memory: It's time for a change of direction. *Journal of Applied Research in Memory and Cognition* **2**, 222–234 (2013).

9. D. Stawarczyk, M. A. Bezdek, J. M. Zacks, Constructing event representations: The role of the midline default network core. *Topics in Cognitive Science* (Submitted).

10. C. N. Wahlheim, J. M. Zacks, Memory guides the processing of event changes for older and younger adults. *Journal of Experimental Psychology: General* **148**, 30–50 (2019).

11. G. A. Radvansky, J. M. Zacks, *Event Cognition* (Oxford University Press, 2014).

12. C. A. Kurby, J. M. Zacks, Segmentation in the perception and memory of events. *Trends Cogn Sci* **12**, 72–9 (2008).

13. J. M. Zacks, C. A. Kurby, M. L. Eisenberg, N. Haroutunian, Prediction Error Associated with the Perceptual Segmentation of Naturalistic Events. *Journal of Cognitive Neuroscience* **23**, 4057–4066 (2011).

14. D. L. Hintzman, Research Strategy in the Study of Memory: Fads, Fallacies, and the Search for the "Coordinates of Truth." *Perspect Psychol Sci* **6**, 253–271 (2011).

15. M. C. Anderson, J. H. Neely, "Interference and inhibition in memory retrieval." in *Memory*, 2nd Ed., E. L. Bjork, R. A. Bjork, Eds. (Academic Press, 1996), pp. 237–313.

16. D. A. Balota, P. O. Dolan, J. M. Duchek, "Memory changes in healthy older adults" in *The Oxford Handbook of Memory*, (Oxford University Press, 2000), pp. 395–409.

17. R. T. Zacks, L. Hasher, K. Z. H. Li, "Human memory" in *The Handbook of Aging and Cognition, 2nd Ed*, (Lawrence Erlbaum Associates Publishers, 2000), pp. 293–357.

18. J. M. Zacks, N. K. Speer, J. M. Vettel, L. L. Jacoby, Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and Aging* **21**, 466–482 (2006).

19. C. A. Kurby, J. M. Zacks, Age differences in the perception of hierarchical structure in events. *Mem Cognit* **39**, 75–91 (2011).

20. C. N. Wahlheim, Proactive effects of memory in young and older adults: the role of change recollection. *Mem Cognit* **42**, 950–64 (2014).

21. C. M. Reese, K. E. Cherry, L. E. Norris, Practical Memory Concerns of Older Adults. *Journal of Clinical Geropsychology* **5**, 231–244 (1999).

22. C. Lustig, T. Jantz, Questions of age differences in interference control: When and how, not if? *Brain Research* https:/doi.org/10.1016/j.brainres.2014.10.024.

23. B. Levine, E. Svoboda, J. F. Hay, G. Winocur, M. Moscovitch, Aging and autobiographical memory: dissociating episodic from semantic retrieval. *Psychol Aging* **17**, 677–689 (2002).

24. T. A. Allen, A. M. Morris, S. M. Stark, N. J. Fortin, C. E. L. Stark, Memory for sequences of events impaired in typical aging. *Learn. Mem.* **22**, 138–148 (2015).

25. P. L. St Jacques, D. Montgomery, D. L. Schacter, Modifying memory for a museum tour in older adults: Reactivation-related updating that enhances and distorts memory is reduced in ageing. *Memory* **23**, 876–887 (2015).

26. N. Kriegeskorte, M. Mur, P. A. Bandettini, Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2** (2008).

27. J. F. Danker, J. R. Anderson, The Ghosts of Brain States Past: Remembering Reactivates the Brain Regions Engaged During Encoding. *Psychological Bulletin* **136**, 87–102 (2010).

28. M. E. Wheeler, S. E. Petersen, R. L. Buckner, Memory's echo: vivid remembering reactivates sensory-specific cortex.[Erratum appears in Proc Natl Acad Sci U S A. 2004 Apr 6;101(14):5181]. *Proc Natl Acad Sci U S A* **97**, 11125–9 (2000).

29. A. M. Gordon, J. Rissman, R. Kiani, A. D. Wagner, Cortical reinstatement mediates the relationship between content-specific encoding activity and subsequent recollection decisions. *Cereb Cortex* **24**, 3350–64 (2014).

30. J. Chen, *et al.*, Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience* **20**, 115–125 (2017).

31. C. M. Bird, J. L. Keidel, L. P. Ing, A. J. Horner, N. Burgess, Consolidation of Complex Events via Reinstatement in Posterior Cingulate Cortex. *J Neurosci* **35**, 14426–34 (2015).

32. C. S. H. Oedekoven, J. L. Keidel, S. C. Berens, C. M. Bird, Reinstatement of memory representations for lifelike events over the course of a week. *Scientific Reports* **7**, 14305 (2017).

33. M. E. Raichle, *et al.*, A default mode of brain function. *Proc Natl Acad Sci U S A* **98**, 676–82 (2001).

34. D. A. Gusnard, M. E. Raichle, Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci* **2**, 685–94 (2001).

35. C. Ranganath, M. Ritchey, Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience* **13**, 713–726 (2012).

36. M. Ritchey, M. E. Montchal, A. P. Yonelinas, C. Ranganath, Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. *elife* **4**, e05025 (2015).

37. A. W. Gilmore, S. M. Nelson, K. B. McDermott, The Contextual Association Network Activates More for Remembered than for Imagined Events. *Cereb Cortex* **26**, 611–7 (2016).

38. A. W. Gilmore, S. M. Nelson, H.-Y. Chen, K. B. McDermott, Task-related and resting-state fMRI identify distinct networks that preferentially support remembering the past and imagining the future. *Neuropsychologia* **110**, 180–189 (2018).

39. M. Bar, The proactive brain: memory for predictions. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **364**, 1235–1243 (2009).

40. D. Stawarczyk, O. Jeunehomme, A. D'Argembeau, Differential Contributions of Default and Dorsal Attention Networks to Remembering Thoughts and External Stimuli From Real-Life Events. *Cereb Cortex* **28**, 4023–4035 (2018).

41. R. K. Olsen, S. N. Moses, L. Riggs, J. D. Ryan, The hippocampus supports multiple cognitive processes through relational binding and comparison. *Front. Hum. Neurosci.* **6** (2012).

42. A. Mayes, D. Montaldi, E. Migo, Associative memory and the medial temporal lobes. *Trends in Cognitive Sciences* **11**, 126–135 (2007).

43. N. Raz, K. M. Rodrigue, Differential aging of the brain: Patterns, cognitive correlates and modifiers. *Neuroscience & Biobehavioral Reviews* **30**, 730–748 (2006).

44. C. Baldassano, *et al.*, Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron* **95**, 709-721.e5 (2017).

45. C. Baldassano, U. Hasson, K. A. Norman, Representation of real-world event schemas during narrative perception. *J. Neurosci.*, 0251–18 (2018).

46. R. Leech, D. J. Sharp, The role of the posterior cingulate cortex in cognition and disease. *Brain* **137**, 12–32 (2014).

47. F. Lin, *et al.*, The cingulate cortex of older adults with excellent memory capacity. *Cortex* **86**, 83–92 (2017).

48. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**, 1–48 (2015).

49. J. Fox, S. Weisberg, *An {R} Companion to Applied Regression*, Second (Sage, 2011).

50. R. Lenth, emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.3.5.1. (2019).

51. D. Lüdecke, sjPlot: Data Visualization for Statistics in Social Science, R package version 2.7.1 (2019).

52. D. Stawarczyk, M. A. Bezdek, J. M. Zacks, Event Representations and Predictive Processing: The Role of the Midline Default Network Core. *Topics in Cognitive Science* (In press) https:/doi.org/10.1111/tops.12450 (September 6, 2019).

53. A. Schaefer, *et al.*, Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex* **28**, 3095–3114 (2018).

54. L. Hoffman, On the Interpretation of Parameters in Multivariate Multilevel Models Across Different Combinations of Model Specification and Estimation. *Advances in Methods and Practices in Psychological Science* **2**, 288–311 (2019).

55. M. D. Fox, *et al.*, The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *PNAS Proceedings of the National Academy of Sciences of the United States of America* **102**, 9673–9678 (2005).

56. D. Vatansever, D. K. Menon, E. A. Stamatakis, Default mode contributions to automated information processing. *PNAS* **114**, 12821–12826 (2017).

57. A. Ben-Yakov, R. N. Henson, The Hippocampal Film Editor: Sensitivity and Specificity to Event Boundaries in Continuous Experience. *J. Neurosci.* **38**, 10057–10068 (2018).

58. A. Gazzaley, M. D'Esposito, *BOLD functional MRI and cognitivec aging* (Oxford University Press., 2005).

59. D. Vatansever, A. Manktelow, B. J. Sahakian, D. K. Menon, E. A. Stamatakis, Default Mode Network Engagement Beyond Self-Referential Internal Mentation. *Brain Connectivity* **8**, 245–253 (2018).

60. J. J. Allaire, *et al.*, rmarkdown: Dynamic Documents for R. R package version 1.13 (2019).

61. M. F. Folstein, S. E. Folstein, P. R. McHugh, "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189–198 (1975).

62. W. C. Shipley, C. C. Burlingame, A convenient self-administering scale for measuring intellectual impairment in psychotics. *AJP* **97**, 1313–1325 (1941).

63. W. Schneider, A. Eschman, A. Zuccolotto, *E-Prime user's guide* (Psychology Software Tools Inc., 2012).

64. B. Whitcher, V. J. Schmid, A. Thorton, Working with the DICOM and NIfTI Data Standards in R. *Journal of Statistical Software* **44**, 1–29 (2011).

65. M. F. Glasser, *et al.*, The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013).

66. M. H. A. Hendriks, N. Daniels, F. Pegado, H. P. Op de Beeck, The Effect of Spatial Smoothing on Representational Similarity in a Simple Motor Paradigm. *Front. Neurol.* **8** (2017).

67. M. Hanke, *et al.*, PyMVPA: a unifying approach to the analysis of neuroscientific data. *Front. Neuroinform.* **3** (2009).

68. J. R. Landis, G. G. Koch, The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174 (1977).

Figures

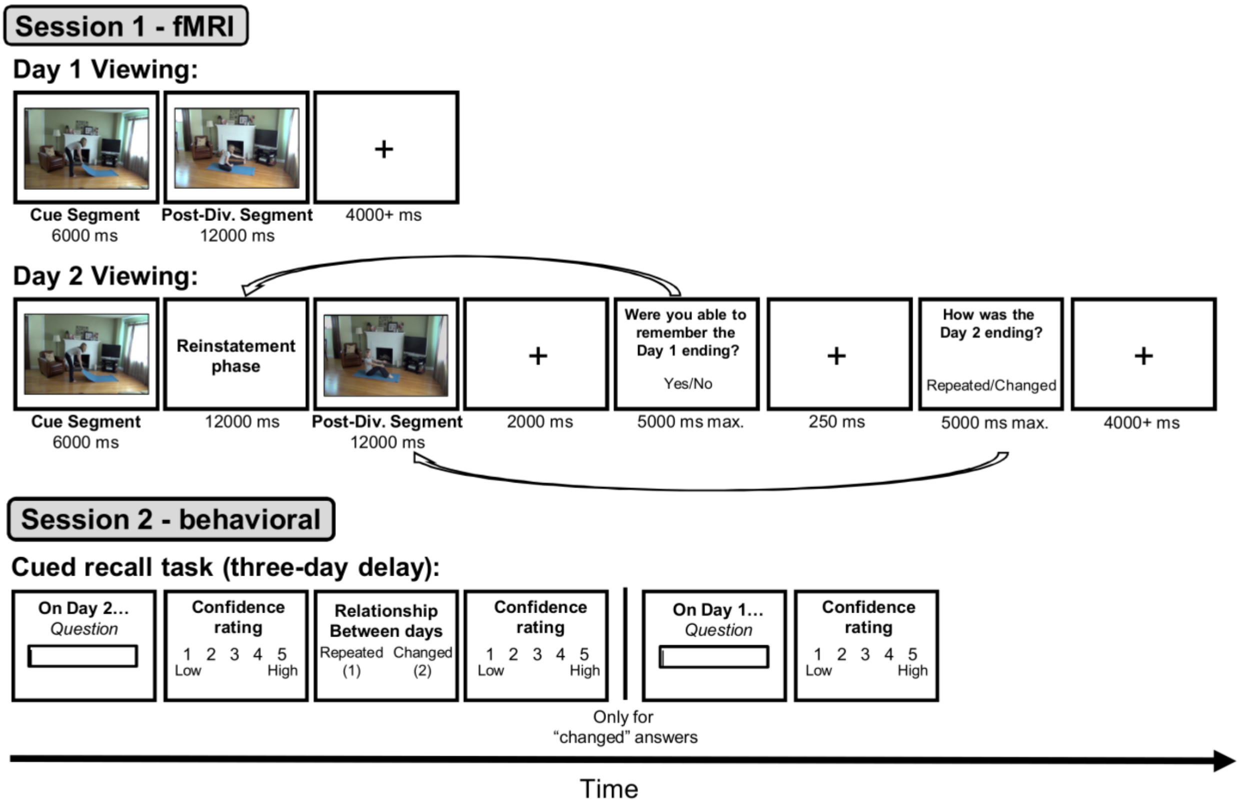


Figure 1. Trial structure of the tasks. During the Day 1 viewing, the transition from the cue to the post-divergence segment of each activity appeared seamlessly. During the Day 2 viewing, the activities were interrupted at the end of the cue segment by a reinstatement phase where participants were asked to mentally replay the corresponding post-divergence segment from the Day 1 viewing. All the cue segments during Day 2 were identical to those of Day 1. Fifteen of the post-divergence segments were repeated exactly from Day 1 while the remaining 30 included a changed feature (as in the illustrated example). The cued recall task took place outside the scanner three days after the first session. The questions about Day 2 and Day 1 were specific to each activity and were always focused on the critical feature that changed between the two versions of each activity (e.g., "What did the actor do on the exercise mat?"). The duration of every step in this task was self-paced. For a more detailed description of the materials and procedure, see the Supporting Information.

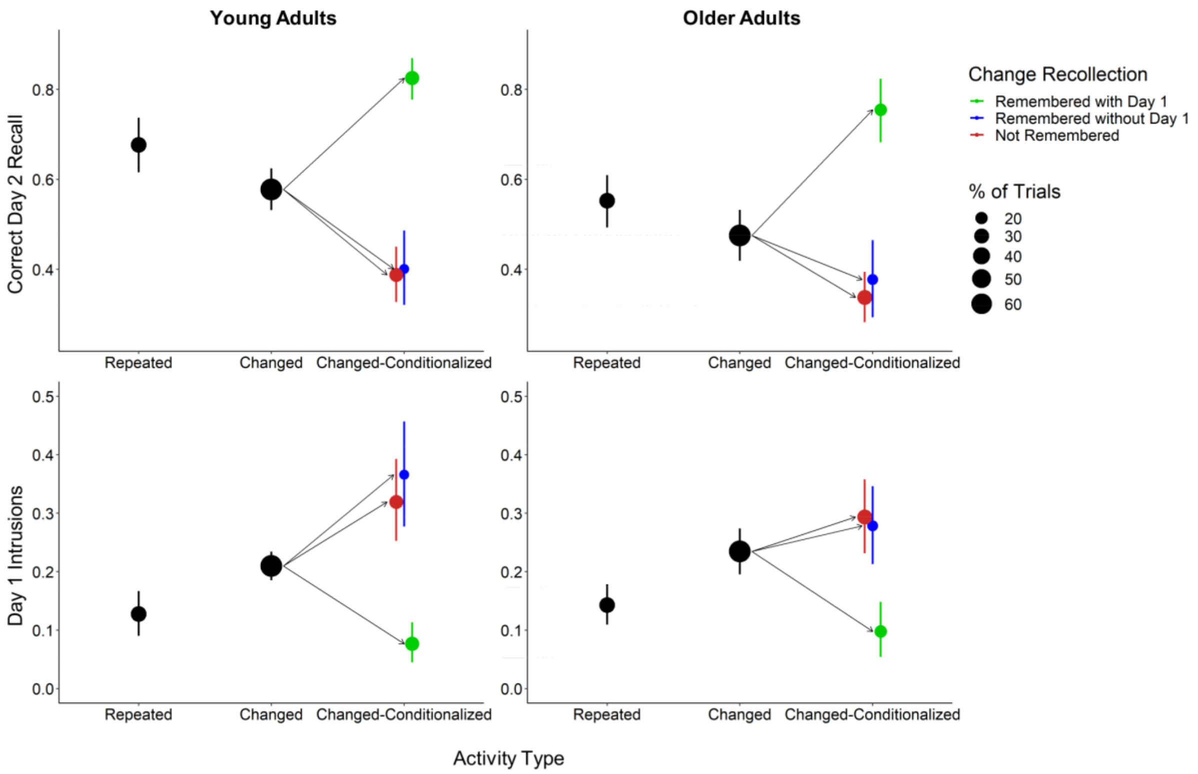


Figure 2. Mean probabilities of correct Day 2 recall and Day 1 intrusions computed over all the trials of each group. Error bars are bootstrap 95% confidence intervals. (Note that intrusions for Repeated activities are cases where participants guessed an ending that they never saw.)

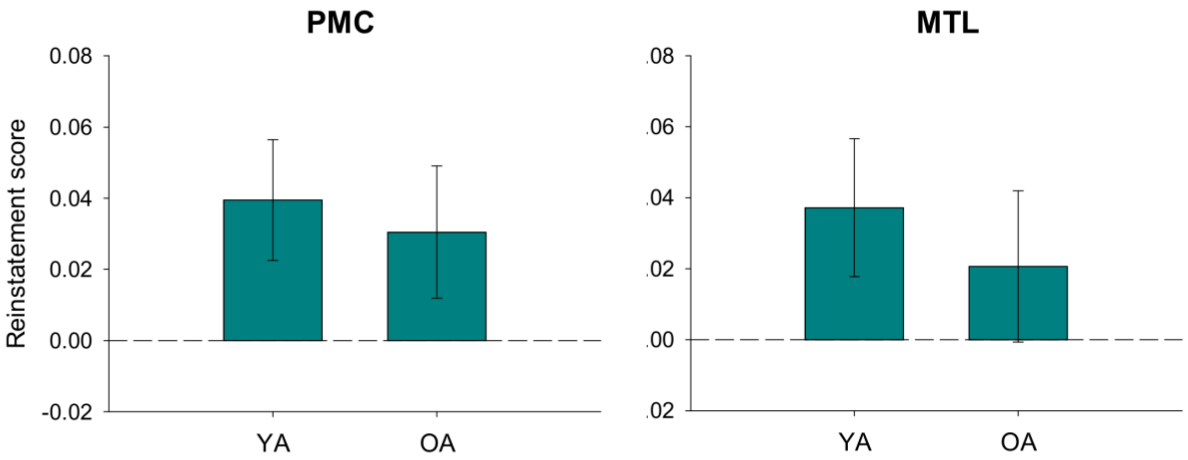


Figure 3. Predicted values for mean reinstatement scores in the MTL and PMC by age group. The error bars represent 95% confidence intervals. YA = Young adults; OA = Older adults.

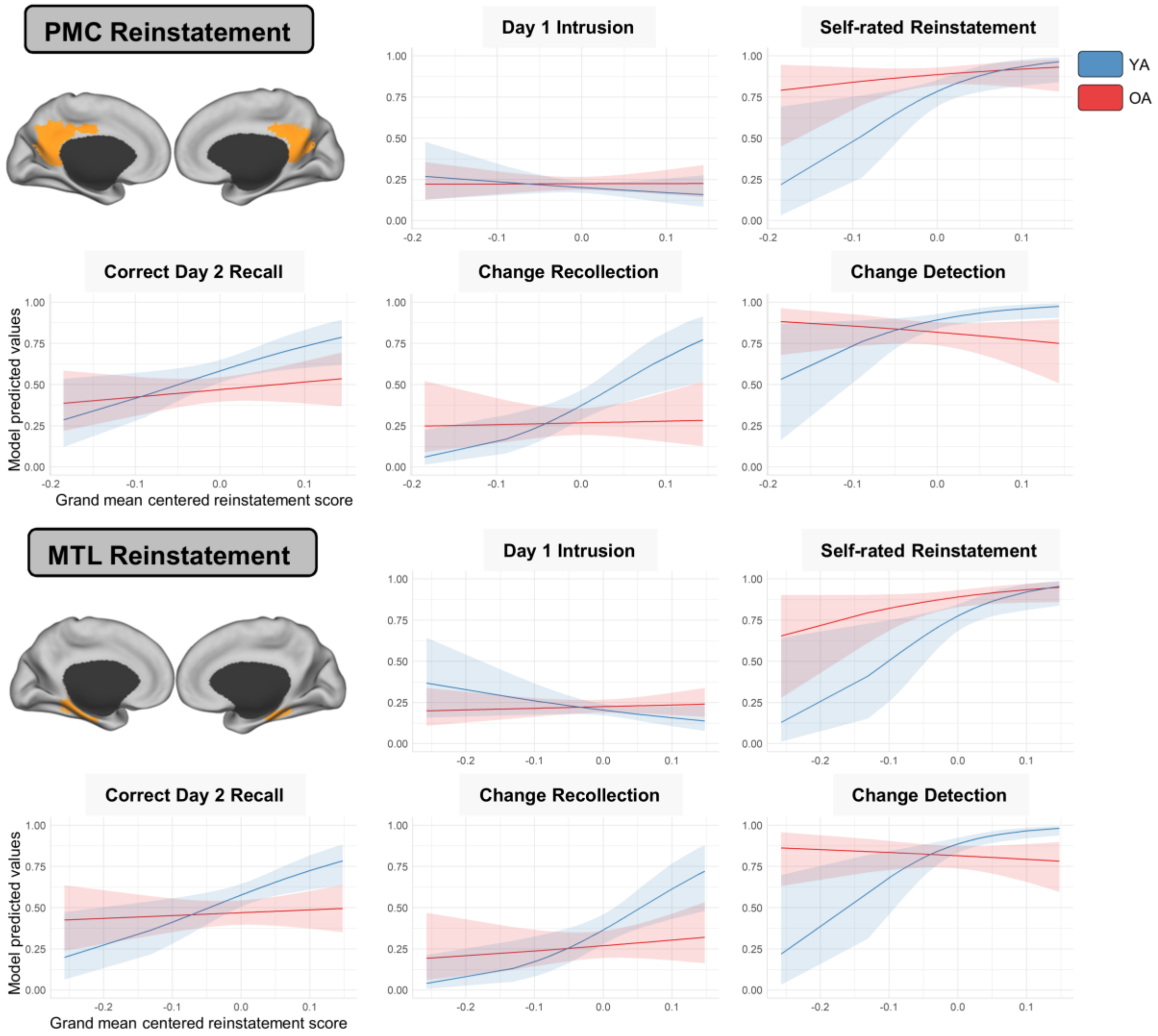


Figure 4. Predicted values and 95% confidence intervals for the association between change comprehension measures and mean PMC/MTL reinstatement scores (between-individual level). YA = Young adults; OA = Older adults.

Table 1: Model predicted values for memory for change and Day 1 recall accuracy for changed activities correctly remembered as changed based on age group.

	Young adults	**Older Adults**
Memory for change	.63 [.54, .70]	.55 [.46, .64]
Day 1 Recall	.69 [.60, .77]	.55 [.44, .65]

Note: 95% confidence intervals are displayed in brackets.

Table 2: Model predicted values for self-reported reinstatement and activity classification accuracy from the Day 2 viewing based on activity type and age group.

		Activity Type	
Dependent variable	**Age group**	**Repeated**	**Changed**
Self-reported reinstatement	Young	.80 [.71, .87]	.79 [.70, .86]
	Older	.89 [.83, .94]	.88 [.81, .93]
Activity Classifications	Young	.87 [.81, .91]	.88 [.83, .91]
	Older	.79 [.71, .85]	.80 [.73, .85]

Note: 95% confidence intervals are displayed in brackets.

Table 3: Model predicted values for change detection accuracy based on self-reported reinstatement accuracy and age group.

	Self-reported reinstatement	
Age group	**Successful**	**Unsuccessful**
Young	.93 [.89, .95]	.79 [.69, .86]
Older	.82 [.75, .88]	.79 [.67, .86]

Note: 95% confidence intervals are displayed in brackets.

Aging and the encoding of changes in events:
The role of neural activity pattern reinstatement

Supporting Information

David Stawarczyk[1, 2], Christopher N. Wahlheim[4], Joset A. Etzel[1], Abraham Z. Snyder[3,5], Jeffrey M. Zacks[1,3]

[1] Department of Psychological & Brain Sciences, Washington University in St. Louis, St. Louis, MO 63105, USA.
[2] Department of Psychology, Psychology and Neuroscience of Cognition Research Unit, University of Liège, 4000 Liège, Belgium.
[3] Mallinckrodt Institute of Radiology, Washington University in St. Louis, St. Louis, MO 63110, USA.
[4] Department of Psychology, University of North Carolina at Greensboro, Greensboro, NC 27412, USA.
[5] Department of Neurology, Washington University in St. Louis, St. Louis, MO 63105, USA.

Correspondence concerning this article should be addressed to David Stawarczyk, Department of Psychology, University of Liège, Quartier Agora, place des Orateurs 1, 4000 Liège 1, Belgium. E-mail: d.stawarczyk@uliege.be. Tel.: 0032 4 3665979.

Table of Contents

1. Supplementary material and Methods 40
2. Supplementary results and discussion 45
3. Recognition task 52
4. References..56
5. Figures 57
6. Tables..70

1. Supplementary material and Methods

Participants

Participants were recruited from the Washington University School of Medicine Research Participant Registry, flyers posted on campus, and word of mouth. Potential participants were initially contacted by phone for a prescreening interview. In addition to MRI contraindications, anyone who reported a history of neurological or psychiatric disorders, taking medication that could affect their cognitive functioning, or not having normal or corrected to normal vision and audition was excluded. The two age groups did not significantly differ in years of education [$t(60) = -1.51$, $p = .14$, $d = .38$; $M = 15.85$ years, $SD = 1.71$ for the young adults; $M = 16.57$ years, $SD = 2.56$ for the older adults] or vocabulary score [$t(60) = -1.87$, $p = .07$, $d = .48$; $M = 83.90\%$, $SD = 8.33$ for the young adults; $M = 87.88\%$, $SD = 8.28$ for the older adults] assessed with the Shipley Institute of Living Scale (1). Each participant received $25.00 per hour for their participation in the study. Nine other participants took part in the study but were excluded from the analyses. Two were young adults who interrupted the study during scanning. Technical issues lead to unusable data for two older adults, and five additional older adults were excluded because they did not comply with the task instructions.

Materials

Each of the 45 activities comprised two parts: a *cue* segment that was identical for the A and B versions and lasted 6 s followed by a *post-divergence* segment that included the changed feature and lasted 12 s with the last second including a fade to black transition. For some of the activities, the changed feature was an object that the actor interacted with (e.g., pouring a glass of milk or a glass of water). For other activities, the changed feature was the action itself (e.g., doing leg stretches or sit-ups on a yoga mat). In all cases the changed feature was central to the activity performed. The critical manipulation was whether the post-divergence segment was the same in both movies (repeated activities), or whether that segment changed from the first to second movie (changed activities; e.g., the A version in the first movie and the B version in the second movie). There were twice the number of changed as compared to repeated activities (i.e., 30 vs. 15) because the primary aim was to investigate neural correlates of change processing.

Colored square-wave gratings were overlaid on the movies during the post-divergence segment with a spatial frequency of approximately half a cycle per degree of visual angle and an opacity of 60%. The gratings gradually appeared over the movies during the first 1.5 s of the post-divergence segment (see Figure S5). Half of the gratings were red and vertical and half were green and horizontal. Gratings were presented in a fixed pseudo-random sequence such that they could not be identical for more than four consecutive activities. For clips that included a changed feature in the second run, the gratings also changed, whereas the gratings remained the same for activities that were repeated across runs[1].

To counterbalance the assignment of activities to conditions, we created 12 experimental formats, by dividing the 45 activities into 3 groups of 15 activities and rotating them through conditions across participants, such that each participant viewed two groups of changed activities and one group of repeated activities. The sequence of activities was fixed, beginning with the actor waking up and ending with her going to bed. The assignment of activities to experimental conditions (which ending was presented on Day 1, and whether the item was repeated or changed) was fixed in a pseudo-random sequence such that each third of

---

[1] The aim of these gratings was to provide an alternative means of analyzing pattern reinstatement that allowed for repeated presentations stimulus features as a complement to identifying unique event features. Initial analyses indicated that reinstatement of event features was more robust than reinstatement of grating features, so the effects of the gratings were not analyzed further.

the task (15 activities) contained five repeated activities with no more than six consecutive changed activities throughout the task. We also alternated the gratings and whether the A or B version of each activity appeared in the first fictive day movie. All stimulus materials were presented using E-Prime 2 software (Psychology Software Tools, Pittsburgh, PA, USA; 3).

Procedure and task description

Participants completed the task in two sessions separated by three days (**Figure 1**). During the first session, participants watched both movies while lying in the scanner. Before the first functional run, we informed participants that they would watch a movie of an actor performing a series of daily activities throughout the course of her day. We instructed participants to pay attention to the actions and the objects she interacted with and told them that the first movie would be referred to as "Day 1." To encourage attention during encoding of post-divergence segments, we suggested that participants associate the movie content with the colored gratings that we told them would also appear a few seconds after the onset of each activity. In the Day 1 movie, the cue segment of each activity was followed immediately by the post-divergence segment. Following each post-divergence segment, a fixation cross appeared in the middle of the screen for a minimum of 4 s (*M* = 4.6 s, *SD* = 1.0 s) with the onset of the cue segment for the next activity being synchronized with the onset of the next scan. This fixation cross provided a demarcation between events, which was necessary for the subsequent reinstatement task. It also allowed the blood-oxygen-level-dependent (BOLD) response to decrease before the onset of the next activity and permitted us to use the exact same time window across activities with regards to the onsets of the video clips in the representational similarity analyses (RSA; see the fMRI analyses section for details). We presented two example activities before the beginning of the run.

After participants viewed the Day 1 movie, field map images and a high-resolution anatomical image were collected, taking approximately 10 minutes. We then told participants that their next task would be to watch another movie that depicted the same actor performing activities on another fictive day that occurred one week later in her life. We called this second functional run the "Day 2" viewing (see **Figure 1**). We explained to participants that the activities would appear in the same order as in the first movie, but that the clips would stop after a few seconds so that they could mentally replay the upcoming action from the Day 1 movie. Participants were told that during this "reinstatement phase" they should imagine the event ending from the Day 1 movie in as much detail as possible. During the Day 2 viewing, each clip stopped after the cue segment, and a question appeared for 12 s (to match the duration of the Day 1 post-divergence segment) asking participants what happened next during Day 1. After the 12 s reinstatement phase, participants watched the post-divergence segment. For this segment, participants were told to pay specific attention to whether the activity features repeated or changed from Day 1. After each post-divergence segment, a 2 s fixation cross followed by two successive questions appeared. The first question asked whether participants thought that they remembered the Day 1 post-divergence segment during the reinstatement phase. The second question asked whether the action in the post-divergence segment was repeated or changed. Participants responded "yes" or "no" to the first question and "repeated" or "changed" to the second question using a button box for both questions. The duration of each question was self-paced with a cut-off of 5 ms, and a 250 ms fixation cross appeared between questions. As for the Day 1 movie, a fixation cross appeared at the end of the trial for a minimum of 4 s (*M* = 4.67 s, *SD* = .39) with the onset of the following cue segment being synchronized with the onset of the next scan.

Before the beginning of the task, participants completed two example trials using the same example activities that appeared before the first run. At the end of the second functional run, we collected a second set of field map images and a high-resolution T2-weighted image. We then informed participants that that the session was over. We did not mention that Session

2 would consist of a memory test in order to reduce the likelihood that participants would rehearse the activities during the three-day delay.

During Session 2, outside of the scanner, we first tested participants' memory for the previously-viewed activities using a cued recall task (see **Figure 1**). During the task, the questions appeared in the same order as the activities during each movie. For each activity, we first asked participants to recall features that appeared in the Day 2 clips by typing their response onto the keyboard. For example, for the question, "What did the actor eat for breakfast?", participants might respond "A banana." The tested features were all critical features that varied between the A and B versions of the activities. After each response, we asked participants whether the way that the actor accomplished the activity changed from Day 1 to Day 2. Participants pressed the "1" key to indicate that the activity repeated exactly and the "2" key to indicate that the activity changed. When participants indicated that the activity had changed, they were prompted to type the original Day 1 feature onto the screen. For each of these questions, we asked participants to rate the confidence in their answer on a Likert scale ranging from 1 "Low" to 5 "High" (these confidence ratings are not relevant to the hypotheses investigated here, so we do not discuss them further). All responding during the task was self-paced. Following the cued recall task, we administered a recognition test of Session 1 activities (see below in the Supporting Information). After the recognition test, all participants completed a vocabulary test (1), and older adults completed the MMSE (4) last.

fMRI data acquisition

All fMRI data were collected using a 3 Tesla scanner (Magnetom Prisma, Siemens Erlangen, Germany) equipped with a 64-channel receiver head coil. Head movement was minimized with foam padding. Anatomical imaging including a multi-echo, T1-weighted MP-RAGE scan (TR = 2500 ms, TE = 1.81, 3.6, 5.39, and 7.18 ms, FOV 256 × 240 mm, matrix 320 × 300 × 208, voxel size 0.8 × 0.8 × 0.8 mm) and a high-resolution T2-weighted SPACE scan (TR = 3200 ms, TE = 564 ms, FOV 256 × 240 mm, matrix 320 × 300 × 208, voxel size 0.8 × 0.8 × 0.8 mm). BOLD fMRI was acquired using a T2*-weighted, multi-band accelerated EPI pulse sequence developed at the Center for Magnetic Resonance Research (CMRR) at the University of Minnesota (MB factor = 4, TR = 1330 ms, TE = 38.8 ms, FA 63°, matrix size 110 × 110 × 60, voxel size 2.4 × 2.4 × 2.4 mm, A >> P phase encoding). Whole brain coverage was obtained with 60 2.4 mm slices without in-plane acceleration (iPAT = 0). Two fMRI runs were acquired in each participant. A mean of 782 (*SD* = 3.83, range: 775-800) and 1343.5 (*SD* = 29.78, range: 1285-1420) volumes were acquired in the first and second runs, respectively. Spin echo field maps were acquired after each fMRI run. Stimuli were displayed on a screen positioned at the rear of the scanner, which participants viewed via a mirror mounted on the head coil.

Spatial preprocessing of fMRI data

The functional data were analyzed in $(2\text{mm})_3$ 711-B2 atlas space (5). Atlas transformation was initially computed by composition of affine transforms (fMRI functional volume mean → T2w → T1w → 711-2B space representative target image). The last transformation step was refined by non-linear registration of each individual's T1w to the atlas representative target using the FNIRT module in fsl (6, 7). One step final resampling of the functional data in atlas space combined retrospective head motion correction, magnetization inhomogeneity distortion correction via topup (8) and non-linear atlas transformation.

2. Supplementary results and discussion

Performance in the cued recall task as a function of self-reported reinstatement and change detection accuracy

We examined the associations of self-reported reinstatement accuracy and change detection with memory measures for changed activities during cued recall (see Figure S6). We used separate 2 (Reinstatement: Successful vs. Unsuccessful) × 2 (Age: Young vs. Older) models for each memory measure (i.e., Day 2 recall, Day 1 intrusions, and Change Recollection). The model for reinstatement accuracy and Day 2 recall (top left panel) indicated that self-reported successful reinstatement was associated with more accurate Day 2 recalls for changed activities [$\chi^2(1) = 12.49$, $p < .001$] and this effect did not interact with Age [$\chi^2(1) = 0.14$, $p = .71$]. The same analysis performed on Day 1 intrusions (middle left panel) revealed a significant Reinstatement × Age interaction [$\chi^2(1) = 4.80$, $p = .03$]. Post-hoc tests revealed that successful reinstatement was associated with fewer intrusions for young adults (z ratio = 2.01, $p = .04$) but not for older adults (z ratio = -1.23, $p = .22$). In addition, successful reinstatement was associated with greater change recollection for both age groups (bottom left panel) [$\chi^2(1) = 61.93$, $p < .001$] and the Reinstatement × Age interaction was not significant [$\chi^2(1) = 0.05$, $p = .82$]. The same models including change detection accuracy as the fixed effect indicated that change detection accuracy was not associated with Day 2 recall (top right panel) [$\chi^2(1) = 3.44$, $p = .07$], nor with Day 1 intrusions (middle right panel) [$\chi^2(1) = 0.01$, $p = .90$], but it was associated with greater change recollection (bottom right panel) [$\chi^2(1) = 87.85$, $p < .001$]. There were no significant interactions with Age (*largest* $\chi^2(1) = 0.74$, $p = .39$).

Finally, because both self-reported reinstatement and change recollection were positively associated with Day 2 recall and were also related to each other, we assessed the EMRC proposal that reinstatement facilitates change recollection by enabling encoding of configural representations when change is initially detected. To do so, we examined whether the effect of self-reported reinstatement remained significant after entering change recollection accuracy in the model as additional fixed effect (see Figure S7 for the Day 2 recall accuracy values predicted by Reinstatement and Change Recollection controlling for the effects of the other variable). A 2 (Reinstatement: Successful vs. Unsuccessful) × 2 (Change Recollection: Recollected vs. Not recollected) × 2 (Age: Young vs. Older) model with Day 2 recall for changed activities as the dependent variable revealed no significant effect of Reinstatement

[$\chi 2(1) = 0.49$, $p = .48$] but the effect of Change Recollection was significant [$\chi 2(1) = 219.95$, $p < .001$]. There were no significant interactions (all *p*s > .05).

RSA results, effects of individual parcels

Analyses of the relations between the mean PMC and MTL reinstatement scores revealed several significant effects (see main text). However, it is unclear whether (1) these effects were driven by specific parcels above and beyond the effects of the other parcels in each of the two areas and whether (2) some significant effects might have been obscured by different parcels within the PMC or MTL having opposite effects on memory performance. To investigate these possibilities, we used logistic linear models with between- and within-individual reinstatement scores but added the individual scores of each parcel as fixed effects rather than a single mean value representing the average of all the parcels. We then used the anova function of the lme4 package to determine whether adding these single parcel scores relative to a baseline model comprising only the random effects improved model fits. We then examined whether these model fits were further improved by adding the interactions terms with Age.

First, regarding Day 2 recall accuracy for changed activities, adding the between-individual reinstatement scores for the PMC parcels did not improve model fit and neither did the interaction terms with Age (all *p*s > .05). These results indicate that the significant effect of the mean PMC reinstatement score reported in the main manuscript was not driven by specific parcels. At the within-individual level, adding the fixed effects of the parcels improved model fit [$\chi 2(12) = 29.12$, $p = .004$] but it was not the case for the interaction terms with Age [$\chi 2(12) = 11.08$, $p = .52$]. Examination of the parameter estimates for the individual parcels showed that higher reinstatement scores in parcels 291 and 292 were associated with higher Day 2 recalls accuracy [$\chi 2(1) = 7.97$, $p = .004$, for parcel 291, $\chi 2(1) = 7.67$, $p = .006$, for parcel 292], whereas parcels 141 and 142 showed the opposite effect [$\chi 2(1) = 3.87$, $p = .049$, for parcel

141, $\chi 2(1) = 6.20$, $p = .01$, for parcel 142]. For the MTL parcels, adding the between-individual reinstatement scores improved model fit [$\chi 2(6) = 13.87$, $p = .03$], but the interactions with Age did not [$\chi 2(6) = 7.22$, $p = .30$]. Examination of the parameter estimates for the individual parcels showed that higher reinstatement in parcel 143 was associated with worse Day 2 recall accuracy [$\chi 2(1) = 4.76$, $p = .03$], but this effect did not remain significant when the fixed effect of Age was included in the model [$\chi 2(1) = 3.72$, $p = .05$]. No significant effect was found at the within-participant level for the MTL parcels (all *p*s > .05).

Next, we performed the same analyses for Day 1 intrusions during the Day 2 recall of changed activities. As with correct Day 2 recalls, adding reinstatement scores for the PMC parcels did not improve model fit at the between-individual level and neither did the interaction with age (all *p*s > .05). However, at the within-participant level, adding the fixed effects of the parcels improved model fit [$\chi 2(12) = 24.68$, $p = .02$], but the interaction terms with Age did not [$\chi 2(12) = 8.55$, $p = .74$]. Examination of the parameter estimates for the individual parcels showed that higher reinstatement scores in parcels 142 and 277 were associated with more Day 1 intrusions [$\chi 2(1) = 6.95$, $p = .008$ for parcel 142, $\chi 2(1) = 4.29$, $p = .04$ for parcel 277] whereas parcels 115 and 292 showed the opposite effect [$\chi 2(1) = 7.62$, $p = .006$ for parcel 115, $\chi 2(1) = 5.93$, $p = .01$ for parcel 292]. No effects were significant for the MTL parcels at either the within- or between-individual level (all *p*s > .05).

Third, regarding Change recollection for changed activities, adding the between-individual reinstatement scores for the PMC parcels did not improve model fit [$\chi 2(12) = 17.03$, $p = .15$], but adding the interaction terms with Age did [$\chi 2(12) = 22.99$, $p = .03$]. Further analyses showed that individual parcel reinstatement scores improved model fit for the young [$\chi 2(12) = 24.49$, $p = .02$] but not the older adults [$\chi 2(12) = 14.66$, $p = .26$]. For the young adults, higher reinstatement score in parcels 115 [$\chi 2(1) = 18.55$, $p < .001$] and 292 [$\chi 2(1) = 5.41$, $p = .02$] were associated with higher change recollection accuracy. No significant effect was found at the within-individual level for the PMC parcels and neither the between- nor within-individual level analyses revealed significant effects for the MTL parcels (all *p*s > .05).

Fourth, for self-reported reinstatement accuracy in the scanner, adding the between-individual reinstatement scores for the PMC parcels did not significantly improve model fit [$\chi 2(12) = 20.71$, $p = .05$], but adding the interaction terms with Age did [$\chi 2(12) = 31.40$, $p = .002$]. Further analyses showed that the individual parcel scores improved model fit for both the young [$\chi 2(12) = 28.40$, $p = .005$] and older adults [$\chi 2(12) = 23.66$, $p = .02$]. However, reinstatement scores in different parcels for each age group were associated with the dependent variable: for the young adults, higher reinstatement in parcels 115 [$\chi 2(1) = 19.83$, $p < .001$] and 117 [$\chi 2(1) = 5.30$, $p = .02$] was associated with higher self-reported reinstatement accuracy whereas it was the opposite for parcel 114 [$\chi 2(1) = 5.83$, $p = .02$]. For the older adults, higher reinstatement scores in parcels 142 [$\chi 2(1) = 10.42$, $p = .001$] and 276 [$\chi 2(1) = 10.48$, $p = .001$] was associated with higher self-reported reinstatement accuracy, whereas it was the opposite for parcel 117 [$\chi 2(1) = 5.12$, $p = .02$]. No significant effect was found at the within- participant level for the PMC parcels and neither the between- nor within-individual level analyses revealed significant effects for the MTL parcels (all *p*s > .05).

Finally, regarding change detection accuracy in the scanner, adding the between-individual reinstatement scores for the PMC parcels improved model fit [$\chi 2(12) = 21.57$, $p = .04$], as did the interaction terms with Age [$\chi 2(12) = 24.97$, $p = .01$]. Further analyses revealed that adding the parcels improved model fit for the older adults [$\chi 2(12) = 32.23$, $p = .001$] but not the younger adults [$\chi 2(12) = 19.36$, $p = .08$]. Examination of the parameter estimates for the individual parcels in the older adults showed that higher reinstatement scores in parcels 114 [$\chi 2(1) = 5.62$, $p = .02$] and 142 [$\chi 2(1) = 18.72$, $p < .001$] was associated with higher change detection accuracy, whereas lower change detection accuracy was associated with higher reinstatement scores in parcels 117 [$\chi 2(1) = 9.14$, $p = .002$] and 275 [$\chi 2(1) = 16.88$, $p < .001$].

No significant effect was found at the within- participant level for the PMC parcels and neither the between- nor within-individual level analyses revealed significant effects for the MTL parcels (all *p*s > .05).

In summary, these analyses mainly showed that reinstatement scores at the within-individual level gave a more complex picture compared to those at between-individual level: No facilitating effects on change comprehension were found for either the PMC or MTL when examining mean reinstatement scores (see the main manuscript). However, when examining the contribution of individual parcels above and beyond the others parcels in each area, within -individual differences in reinstatement scores for some PMC parcels were associated with better memory for the changed features, whereas the opposite was true for other parcels (e.g., parcel 292 in the right retrosplenial cortex vs. 142 in the left retrosplenial cortex). To the best of our knowledge, these results are the first to reveal that reinstatement in closely located areas within the same cortical region can have opposite effects on subsequent memory performance. Interestingly, reinstatement in some of these parcels (e.g., parcel 142) were both predictive of worse memory for the changed features and a higher probability of committing intrusions by erroneously recalling the original activity feature. These results suggest that reinstatement in these areas did not purely consist of noise and did reflect memory for the activity features of the first viewing, but in a way that did not facilitate their association with subsequently presented changed activity features. Together, these results foster the view that extended cortical areas such as the PMC are composed of distinct subregions (9–11), and indicate that—at the within-individual level—pattern reinstatement in some PMC parcels might not be beneficial but rather hinder memory for changes, possibly by disrupting the creation of configural representations and thus favoring the experience of proactive interference (12).

Conjoint effect of self-reported reinstatement and RSA reinstatement scores on Day 2 recall accuracy for changed activities

Because both self-reported and neural measures of reinstatement were positively associated with higher memory accuracy for the changed Day 2 activity features for young adults, and because neither effects remained significant after controlling for change recollection performance (see main manuscript), we tested whether self-reported reinstatement accuracy and RSA reinstatement scores were significantly associated with Day 2 recall accuracy for changed features above and beyond each other. These models thus included both self-reported reinstatement (Successful vs. Unsuccessful) and mean reinstatement scores at the between-individual level for either the PMC or MTL as fixed effects (see **Figure S8** for the predicted Day 2 recall accuracy values by mean reinstatement scores and self-reported reinstatement above and beyond the effect of each other). Both the PMC and MTL reinstatement scores [$\chi 2(1) = 5.37$, $p = .02$, for the PMC; $\chi 2(1) = 6.77$ $p = .009$, for the MTL], and self-reported reinstatement were significant in these models [$\chi 2(1) = 10.09$, $p = .001$, for the PMC model; $\chi 2(1) = 9.81$, $p = .002$, for the MTL model]; neither interaction terms were significant (both *ps* > .05).

These results showed that between-individual differences in neural reinstatement scores and self-reported reinstatement accuracy both independently predicted recall performance for the changed features in the young adults. In addition, self-reported, but not neural reinstatement, predicted correct recognition of the changed features (see below in the recognition task section). Although the effects of these two indices of reinstatement on memory for the Day 2 changed feature was fully explained by change recollection, these results suggest that these two measures reflect partly distinct but complementary processes. A possibility is that young participants mostly made their reinstatement judgments based on whether they were able to remember the specific features that changed between the two movies, whereas neural reinstatement in the PMC and MTL reflected the retrieval of more abstract features of the event models (e.g., spatio-temporal context within the movies of each activity) but not necessarily the changed features. It might explain why reinstatement accuracy was not associated with performance in the recognition task where the two response alternatives always displayed the correct activity and only differed with regards to the changed features (see **Figure S9**). Another possibly is that, although participants had to make their reinstatement judgements based on what they remembered during the reinstatement phase, their responses might have been influenced by their experience of the post-divergence segments (i.e., a kind of hindsight bias□judging one's performance as more successful after being presented with correct information (13)). Hindsight bias effects are larger in older adults (14), which may also explain why older adults indicated greater success in their self-reported reinstatement judgments than younger adults did. Future studies might limit this issue with more objective behavioral measures of reinstatement accuracy (e.g., verbal reports of memory representations retrieved during the reinstatement phase).

3. Recognition task

Task description

Following the cued recall task, we tested participants' memory for Session 1 activities using a recognition task (**see Figure S9**). We presented the same questions in the same order as in the cued recall test, but we asked participants to respond by choosing which of two pictures appeared in the Day 2 video. Each picture showed a key frame from the A or B version of the activity without gratings. The pictures appeared on opposite sides of the screen, and the version of the activity that was presented on the left or right side was counterbalanced across trials such that no more than three A or B pictures could consecutively appear on the same side of the screen. Participants responded by pressing the "1" key for the left picture and the "2" key for the right picture. After responding, the picture that was not chosen appeared in the center of the screen, and participants were prompted to indicate whether this alternate version of the activity appeared in the Day 1 movie. They responded by pressing the "1" key for "yes" (to indicate that the activity had changed) and the "2" key for "no" (to indicate that the activity was repeated). Again, we asked participants to rate their confidence in the accuracy of their response for each of these two questions on the same Likert scale as in the cued recall task (ranging from 1 "Low" to 5 "High"). As with the cued recall test, all responses were self-paced.

Behavioral Results

As expected and consistent with the results of the cued recall task, **Figure S10** shows that Day 2 recognition accuracy was higher for repeated than changed activities [$\chi^2(1) = 16.29$, $p < .001$] and that younger adults had higher performance than older adults [$\chi^2(1) = 9.03$, $p = .002$] with no significant Activity × Age interaction [$\chi^2(1) = 0.50$, $p = .48$]. In addition, changed activities for which participants reported having seen both versions of the activity (and thus remembered that a feature changed) were associated with higher Day 2 recognition accuracy than changed activities for which participants indicated not having seen the Day 1 version of the activity [$\chi^2(1) = 61.95$, $p < .001$]. This effect was not qualified by a significant interaction with Age [$\chi^2(1) = 0.81$, $p = .37$]. **Figure S11** (left panels) shows that the effect of self-reported reinstatement success on Day 2 recognition accuracy for changed activities was qualified by a significant interaction with Age [$\chi^2(1) = 6.44$, $p = .01$]. Post-hoc tests revealed that self-reported reinstatement success was associated with higher Day 2 recognition accuracy for changed activities for young (z ratio = 2.80, $p = .005$) but not older adults (z ratio = 0.10, $p = .31$). Self-reported successful reinstatements were also associated with a higher probability of reporting having seen both versions of the activity in the recognition task [$\chi^2(1) = 12.76$, $p < .001$]. This effect did not significantly interact with Age [$\chi^2(1) = 0.23$, $p = .63$]. Finally, **Figure S11** (right panels) shows that change detection accuracy was not associated with Day 2 recognition accuracy for changed activities and the interaction with Age was not significant (all *ps* > .05).

However, change detection accuracy was associated with more reports of having seen both versions of the activity for changed activities [$\chi2(1) = 39.22$, $p < .001$] and this effect was not qualified by an interaction with Age [$\chi2(1) = 1.31$, $p = .25$].

RSA results: mean PMC and MTL reinstatement scores

Recognition performance predicted from the between- and within-individual reinstatement terms are presented in **Figure S12.** Regarding Day 2 recognition accuracy for changed activities, analyses revealed no significant effect of the mean PMC or MTL reinstatement scores, either at the between- or within-individual level and no significant interaction with Age (all *ps* > .05). Regarding the tendency to remember change by reporting having seen both versions of the activities for changed activities, there was no effect of mean PMC reinstatement score, either at the between- of within-participant level (all *ps* > .05). However, for the MTL, the interaction terms between Age and the reinstatement score at both the between- and within-individual level were significant [$\chi2(1) = 4.72$, $p = .03$, for the between-individual level; $\chi2(1) = 8.76$, $p = .003$, for the within-individual level]. Further analyses showed that young adults with higher mean MTL reinstatement scores reported having seen both version of the changed activities more often [$\chi2(1) = 5.85$, $p = .02$] but this what not the case for the older adults [$\chi2(1) = 0.08$, $p = .78$]. Surprisingly, at the within-individual level for the young adult group, trials with higher reinstatement scores were associated with a *lower* tendency to report having seen both versions of the activity [$\chi2(1) = 5.67$, $p = .02$], whereas no significant effect was found for older adults [$\chi2(1) = 3.08$, $p = .08$].

RSA results: individual parcel scores

Regarding Day 2 recognition accuracy for changed activities, adding the between-individual reinstatement scores for the PMC parcels did not improve model fit, and neither did the interaction terms with Age (all *p*s > .05). However, at the within-participant level, adding the fixed effects of the parcels improved model fit [$\chi2(12) = 8.89$, $p = .004$], but the interaction terms with Age did not [$\chi2(12) = 9.35$, $p = .67$]. Examination of the parameter estimates for the individual parcel showed that higher reinstatement scores in parcel 291 were associated with higher Day 2 recognition accuracy [$\chi2(1) = 12.28$, $p < .001$], whereas parcel 276 showed the opposite effect [$\chi2(1) = 4.46$, $p = .03$]. No effects were significant for the MTL parcels at either the within- or between-individual level, and no effects were significant for either the PMC or MTL parcels when examining the association between reinstatement scores and reporting having seen both version of the activity (all *ps* > .05).

4. References

1. W. C. Shipley, C. C. Burlingame, A convenient self-administering scale for measuring intellectual impairment in psychotics. *AJP* **97**, 1313–1325 (1941).

2. C. N. Wahlheim, J. M. Zacks, Memory guides the processing of event changes for older and younger adults. *Journal of Experimental Psychology: General* **148**, 30–50 (2019).

3. W. Schneider, A. Eschman, A. Zuccolotto, *E-Prime user's guide* (Psychology Software Tools Inc., 2012).

4. M. F. Folstein, S. E. Folstein, P. R. McHugh, "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189–198 (1975).

5. R. L. Buckner, *et al.*, A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. *NeuroImage* **23**, 724–738 (2004).

6. J. L. R. Andersson, M. Jenkinson, S. Smith, Non-linear registration, aka spatial normalisation. *FMRIB technical report TR07JA2* (2010).

7. M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, S. M. Smith, FSL. *NeuroImage* **62**, 782–790 (2012).

8. J. L. R. Andersson, S. Skare, J. Ashburner, How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* **20**, 870–888 (2003).

9. R. Leech, D. J. Sharp, The role of the posterior cingulate cortex in cognition and disease. *Brain* **137**, 12–32 (2014).

10. A. Schaefer, *et al.*, Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex* **28**, 3095–3114 (2018).

11. R. Leech, R. Braga, D. J. Sharp, Echoes of the brain within the posterior cingulate cortex. *J Neurosci* **32**, 215–22 (2012).

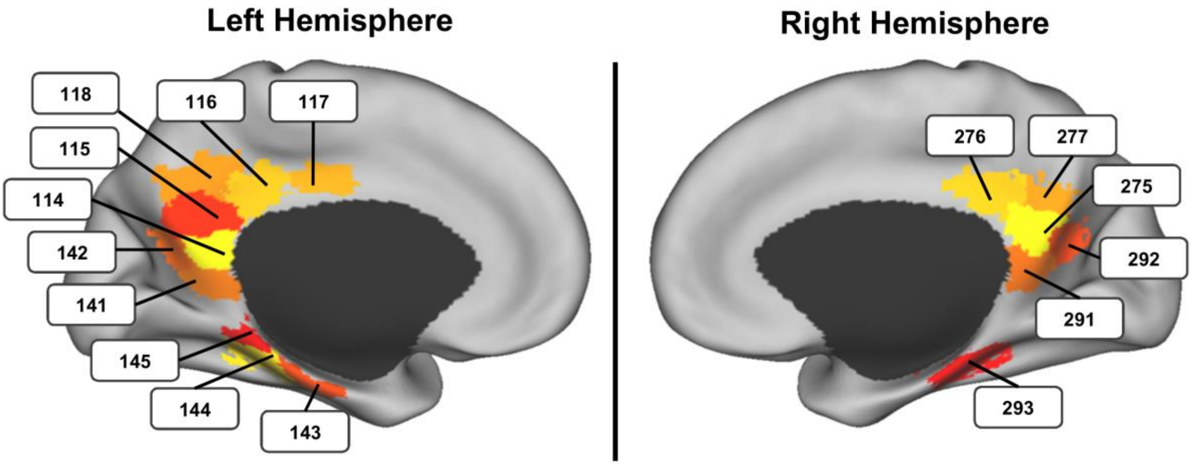
12. M. C. Anderson, J. H. Neely, "Interference and inhibition in memory retrieval." in *Memory*, 2nd Ed., E. L. Bjork, R. A. Bjork, Eds. (Academic Press, 1996), pp. 237–313.

13. R. Hertwig, G. Gigerenzer, U. Hoffrage, The reiteration effect in hindsight bias. *Psychological Review* **104**, 194–202 (1997).

14. U. J. Bayen, E. Erdfelder, J. N. Bearden, J. P. Lozito, The interplay of memory and judgment processes in effects of aging on hindsight bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **32**, 1003–1018 (2006).

5. Figures

Figure S1. Parcels of interest



Note: PMC and MTL parcels of interest of the DN subsystems from the 300 parcels of the 17 networks cortex parcellation map of Schaefer et al. (10). The numbers attributed to each parcel correspond to the labels in the original parcellation. Parcels 143-145 and 293 correspond to the MTL, all other parcels correspond to the PMC.

Figure S2. Reinstatement scores across the full set of parcels

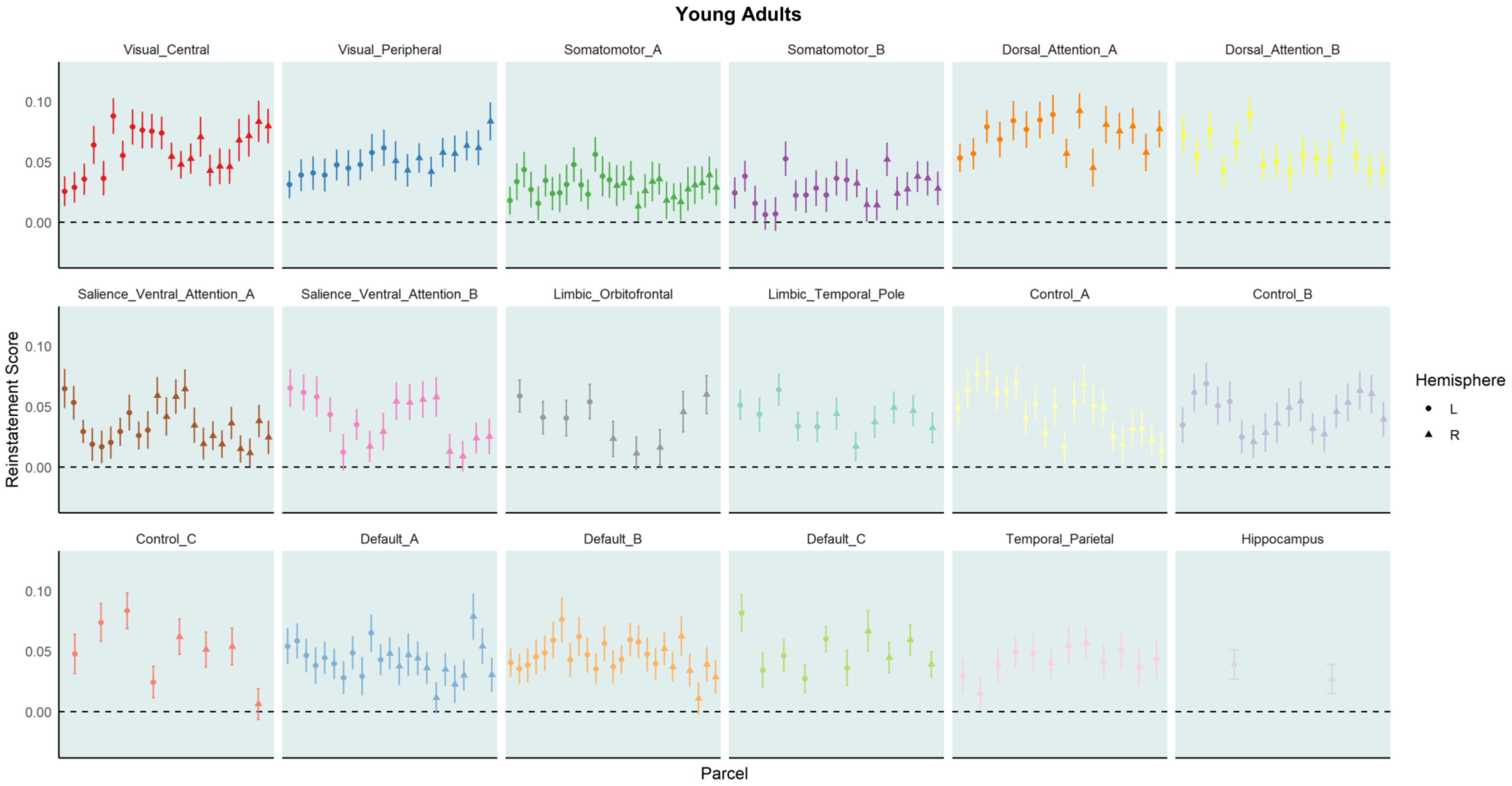
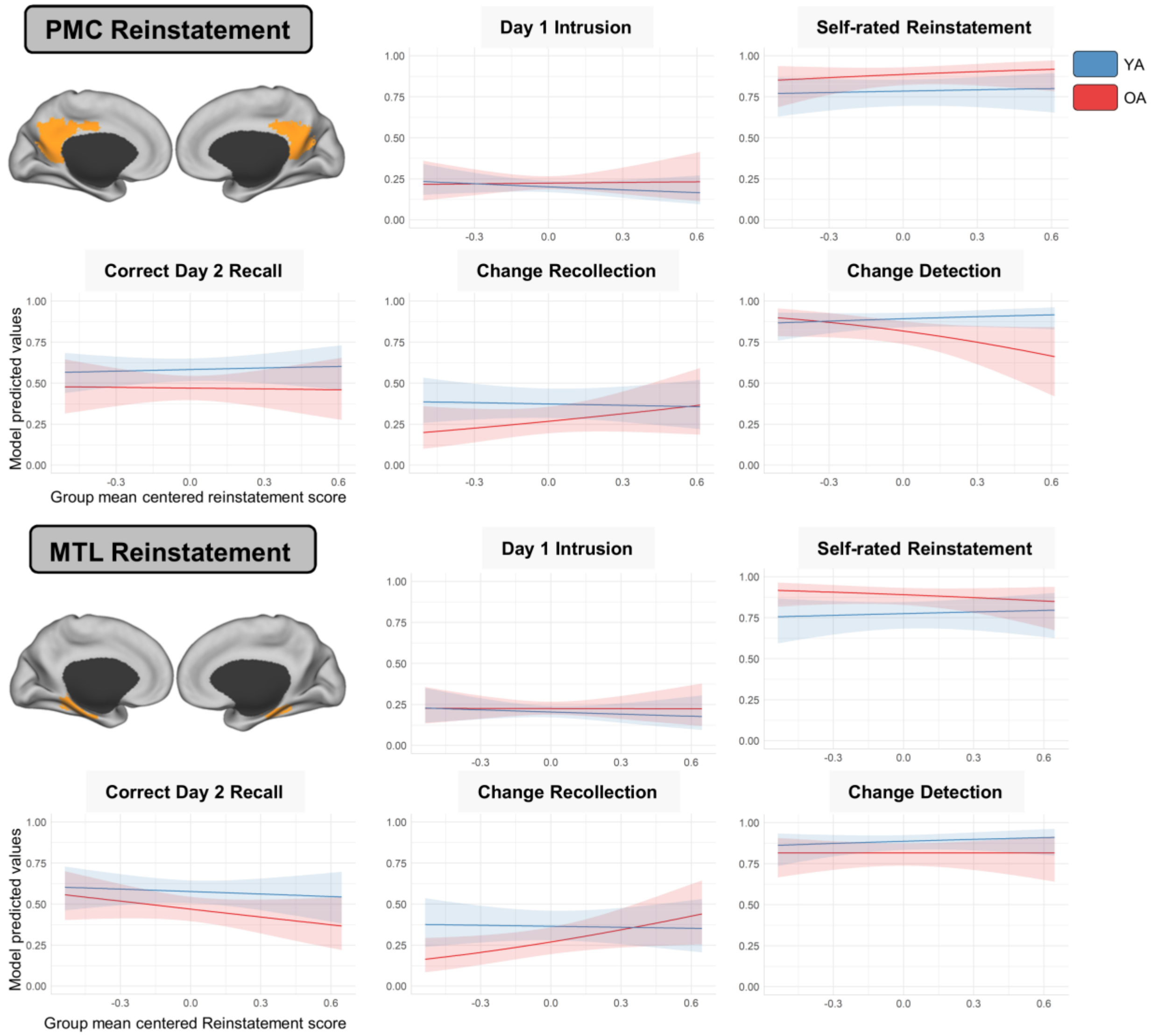


Figure S2. Reinstatement scores (continued)

Note: Mean reinstatement scores and 95% confidence intervals computed over all the trials of each group for the 300 parcels of the 17 network parcellation of Schaefer et al. (1). The parcels and networks are ordered within each hemisphere following the original parcellation numbering.

Figure S3. Association between change comprehension measures in the cued recall task and mean PMC/MTL reinstatement scores (within-individual level).



Note: YA = Young Adults; OA = Older Adults.

Figure S4. Day 2 recall accuracy by mean PMC/MTL reinstatement scores (between-individual level) and change recollection accuracy above and beyond the effect of the other variable for young adults.

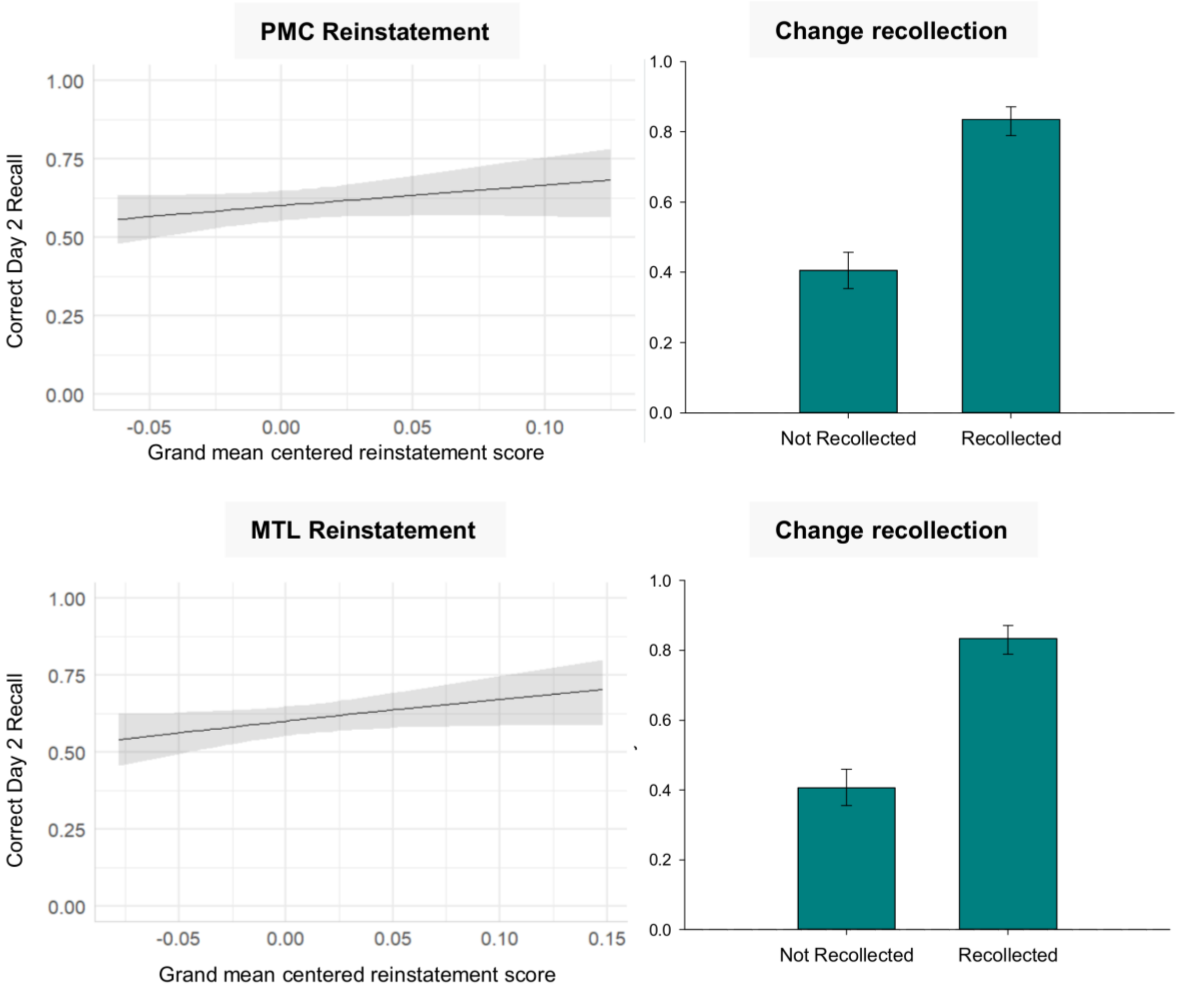
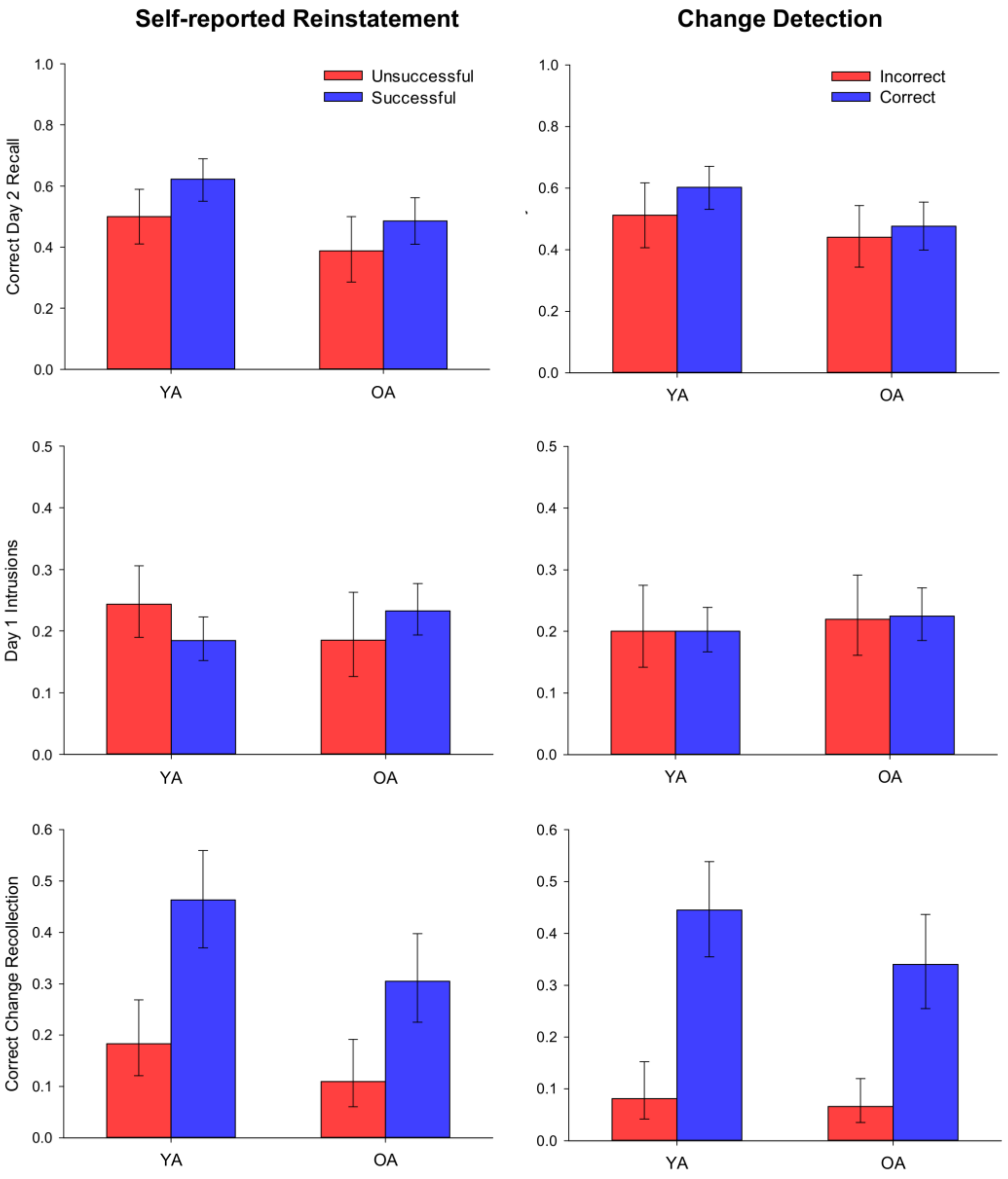


Figure S5. Illustration of the gratings overlayed on the movies during the post-divergence segments of the activities

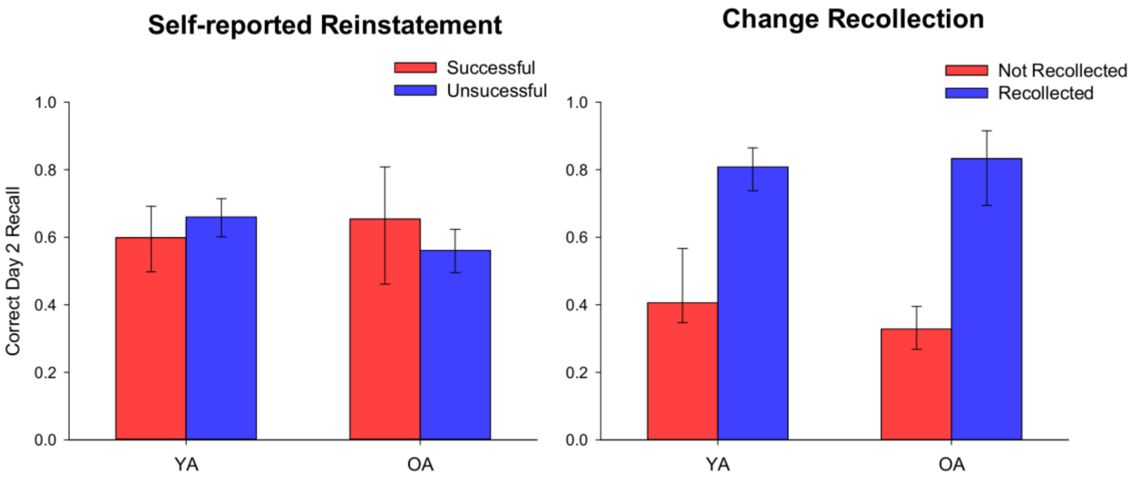
Note: The actor’s face has been obscured in compliance with bioRxiv policy.

Figure S6. Performance in the cued recall task for changed activities for young and older adults as a function of self-reported reinstatement and change detection accuracy.



Note: The predicted values are estimated from the logistic mixed models. Error bars illustrate the 95% confidence interval. YA = Young Adults; OA = Older Adults.

Figure S7. Day 2 recall accuracy by self-reported reinstatement and change recollection accuracy above and beyond the effect of the other variable for young and older adults.



Note: The predicted values are estimated from the logistic mixed models. Error bars illustrate the 95% confidence interval. YA = Young Adults; OA = Older Adults

Figure S8. Day 2 recall accuracy by mean PMC/MTL reinstatement scores (between-individual level) and self-reported reinstatement accuracy above and beyond the effect of the other variable for young adults.

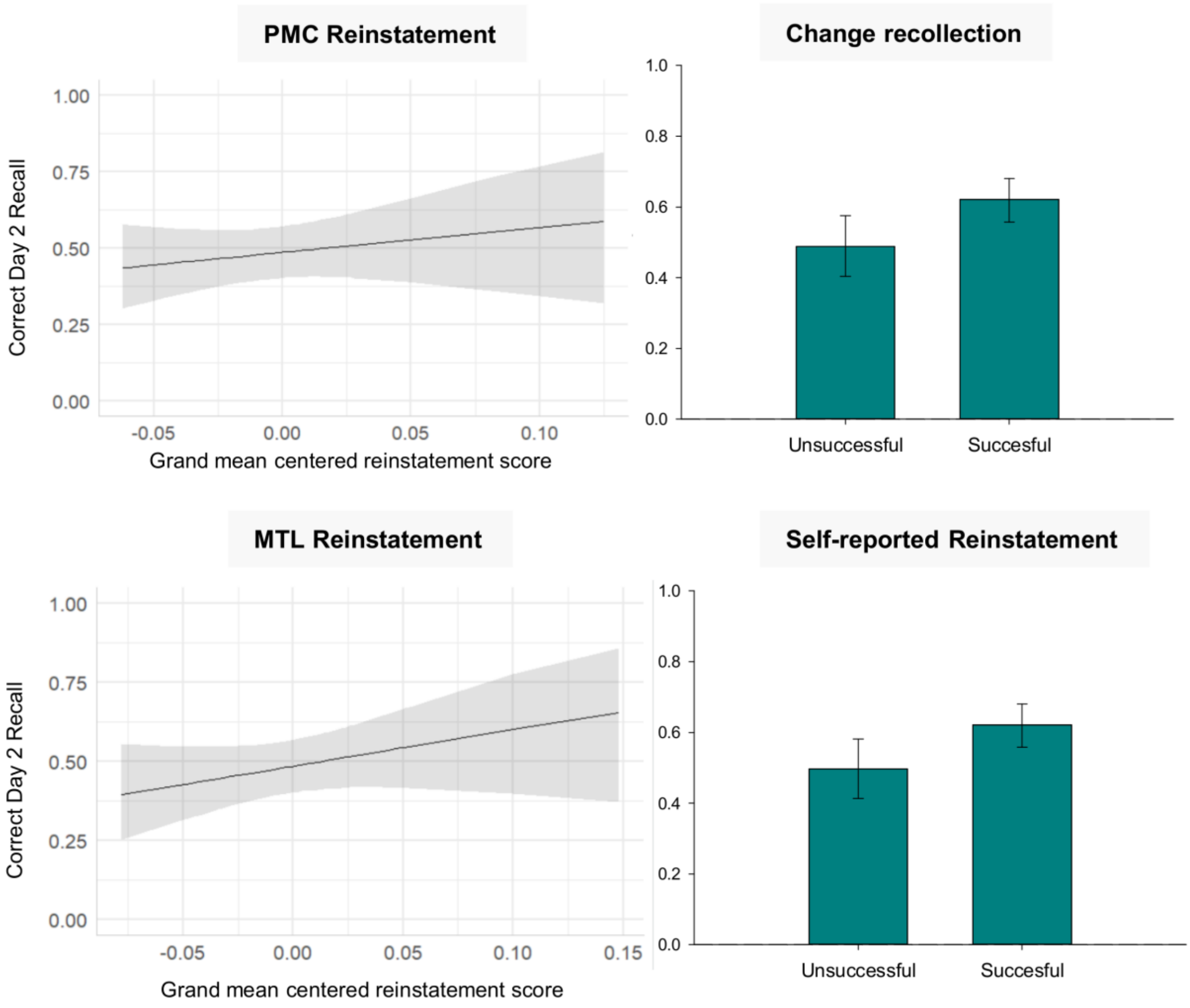
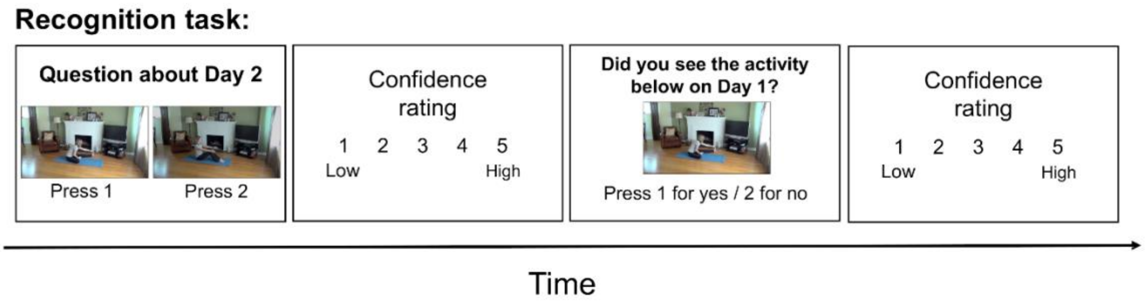
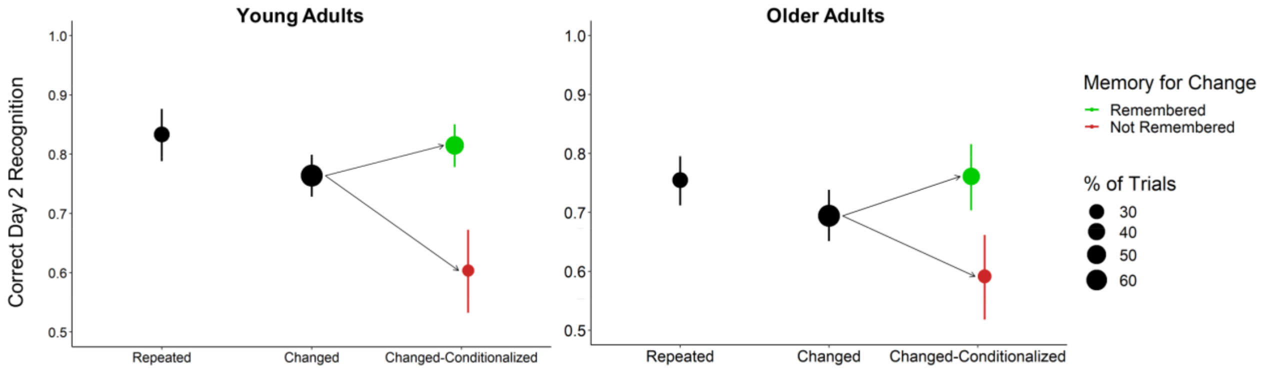


Figure S9. Temporal description of the recognition task.



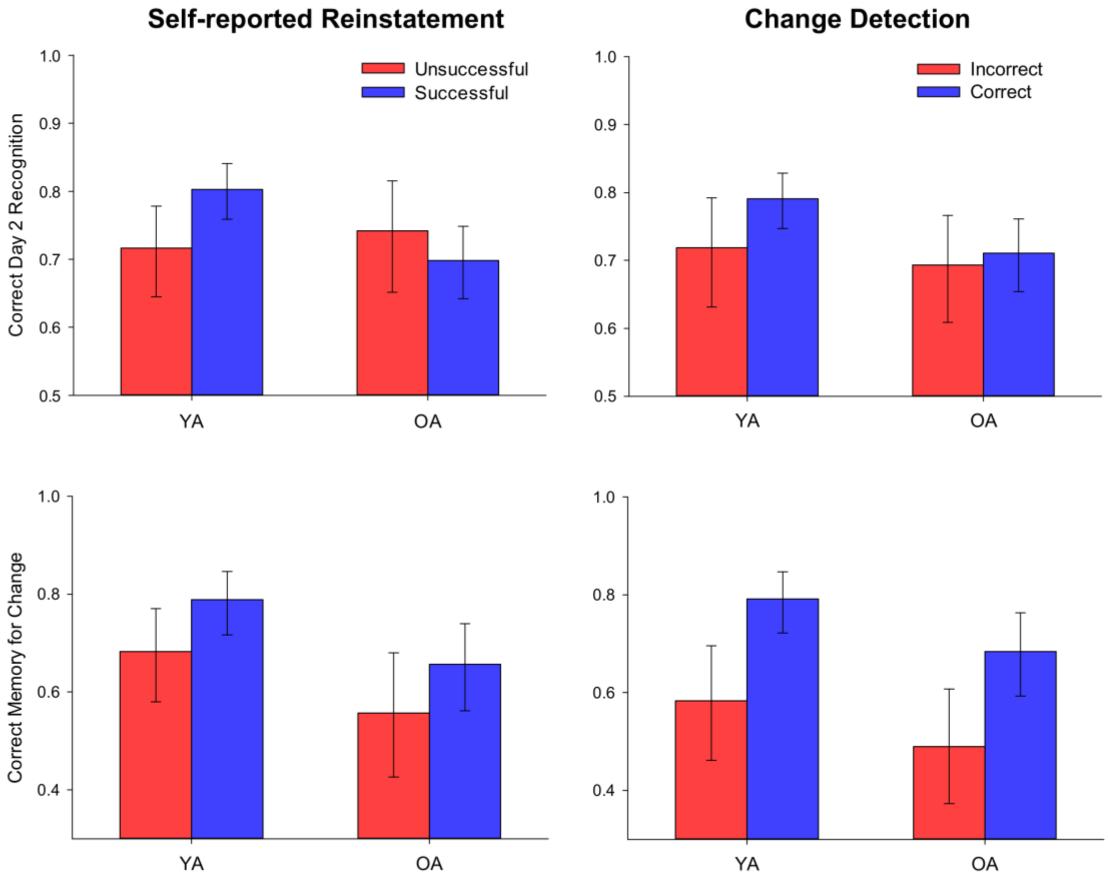
Note: the single picture showed for the Day 1 recognition was always the picture that the participants did not chose in the dual-choice question about Day 2. A "Yes" answer to that question indicated that the participants judged that they saw both versions of the activity and thus that it was changed; a "No" answer indicated that the participant judged that they did not see the alternate version of the activity on Day 1 and thus that is was repeated on both days. The duration of every steps in the task was self-paced.

Figure S10. Probabilities of correct Day 2 recognition for young and older adults. Memory for change is based on the response to the Day 1 recognition question (see Figure S9).



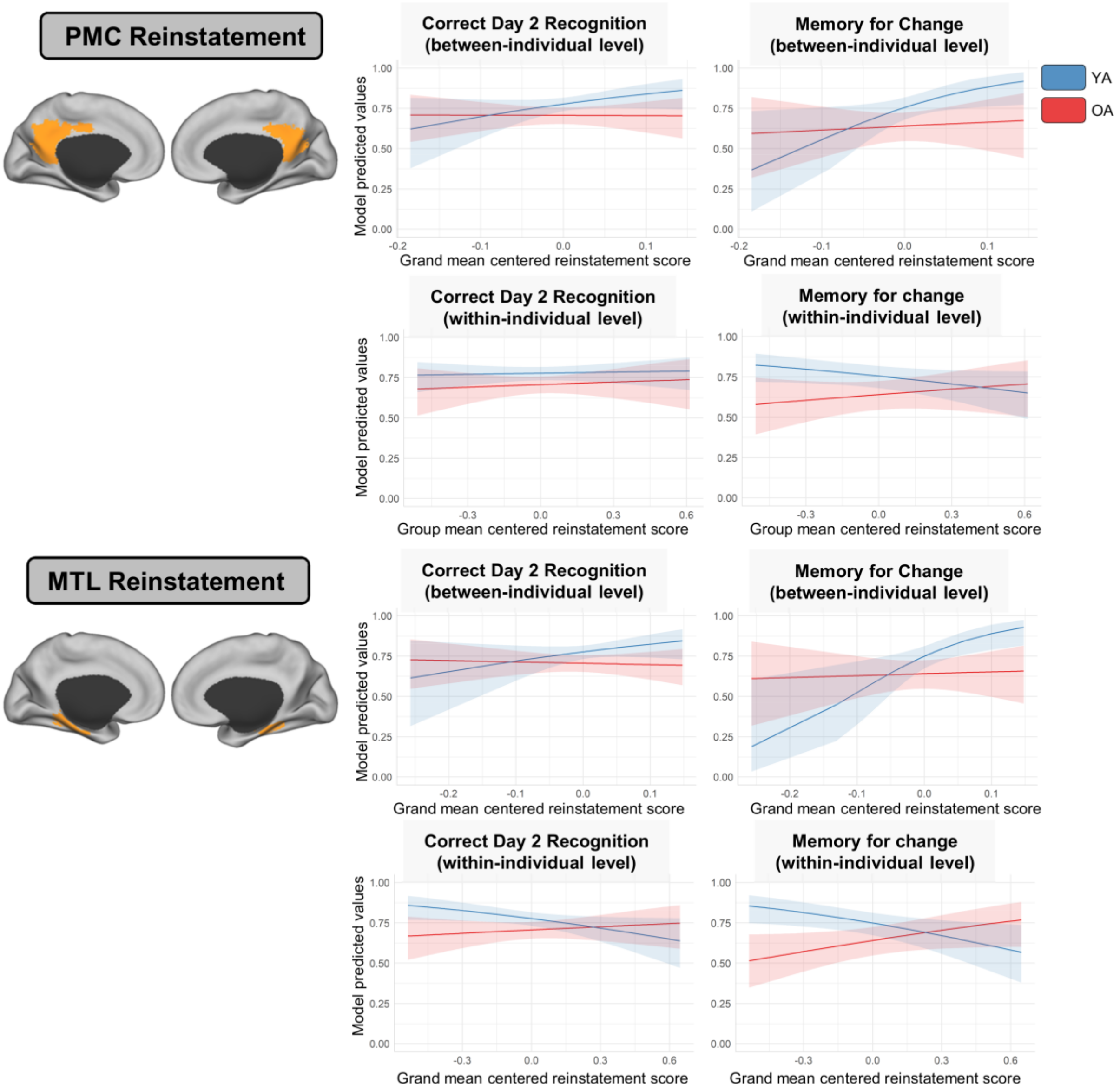
Note: Error bars are bootstrap 95% confidence intervals.

Figure S11. Performance in the recognition task for changed activities for young and older adults as a function of self-reported reinstatement and change detection accuracy.



Note: The predicted values and 95% confidence intervals are estimated from the logistic mixed models. YA = Young Adults; OA = Older Adults.

Figure S12. Association between recognition task performance and mean PMC/MTL reinstatement scores.



Note: YA = Young Adults; OA = Older Adults**.**

6. Tables

Table S1. Comparisons of reinstatement scores for young and older adults for each PMC and MTL parcel.

Parcel	χ^2(1)	*p*-value	Parameter estimates YA [95% CI]	Parameter estimates OA [95% CI]
PMC Left				
p.114	1.17	.28	0.04 [0.02; 0.07]	0.02 [-0.001; 0.05]
p.115	0.32	.57	0.05 [0.02; 0.07]	0.04 [0.01; 0.06]
p.116	0.94	.33	0.04 [0.02; 0.06]	0.02 [-0.0001; 0.05]
p.117	0.03	.87	0.03 [0.004; 0.06]	0.03 [0.004; 0.06]
p.118	0.08	.78	0.05 [0.02; 0.08]	0.05 [0.02; 0.07]
p.141	0.81	.37	0.03 [0.004; 0.06]	0.01 [-0.02; 0.04]
p.142	0.02	.90	0.05 [0.02; 0.07]	0.05 [0.02; 0.07]
PMC Right				
p.275	0.32	.57	0.04 [0.02; 0.05]	0.03 [0.01; 0.05]
p.276	0.01	.90	0.005 [-0.02; 0.03]	0.003 [-0.02; 0.03]
p.277	0.52	.47	0.05 [0.02; 0.07]	0.04 [0.01; 0.06]
p.291	0.89	.34	0.04 [0.02; 0.07]	0.03 [0.002; 0.05]
p.292	0.08	.77	0.07 [0.04; 0.09]	0.06 [0.03; 0.09]
MTL Left				
p.143	0.13	.72	0.02 [-0.004; 0.05]	0.02 [-0.01; 0.05]
p.144	4.47	.03	0.06 [0.04; 0.09]	0.02 [-0.002; 0.05]
p.145	1.32	.25	0.03 [0.003; 0.06]	0.01 [-0.03; 0.04]
Hippo.left	1.19	.27	0.04 [0.02; 0.06]	0.02 [-0.01; 0.04]
MTL Right				
p.293	0.24	.62	0.04 [0.02; 0.06]	0.03 [0.01; 0.06]
Hippo.Right	0.36	.55	0.02 [0.004; 0.05]	0.01 [-0.01; 0.04]

Table S2. Pearson correlation matrix of reinstatement scores.

	114	115	116	117	118	141	142	277	275	276	291	292	143	144	145	Hip.L	293
115	.43																
116	.27	.39															
117	.16	.21	.28														
118	.28	.50	.39	.22													
141	.35	.32	.24	.13	.28												
142	.31	.36	.22	.12	.32	.34											
277	.26	.44	.29	.19	.46	.25	.26										
275	.42	.52	.30	.18	.35	.27	.28	.44									
276	.29	.36	.48	.29	.36	.25	.23	.37	.36								
291	.35	.32	.27	.18	.30	.42	.30	.29	.37	.33							
292	.25	.28	.19	.13	.27	.30	.42	.31	.33	.27	.39						
143	.16	.18	.20	.21	.18	.21	.14	.15	.13	.19	.20	.15					
144	.19	.24	.26	.16	.27	.32	.28	.22	.18	.24	.28	.27	.41				
145	.16	.21	.21	.14	.19	.31	.18	.18	.16	.22	.27	.19	.32	.41			
Hip.L	.18	.20	.18	.18	.20	.22	.18	.16	.17	.19	.20	.17	.42	.47	.28		
293	.20	.25	.21	.18	.22	.25	.23	.22	.22	.29	.34	.27	.34	.37	.26	.29	
Hip.R	.13	.16	.17	.16	.18	.16	.14	.17	.18	.24	.24	.18	.26	.24	.14	.34	.54

Note: The coefficients were obtained by correlating the reinstatement score of each trial across all participants.