

## Cognitive Neuroscience of Aging

# How Multiple Retrievals Affect Neural Reactivation in Young and Older Adults

Marie St-Laurent, PhD<sup>1</sup> and Bradley R. Buchsbaum, PhD<sup>2,3,\*</sup>

<sup>1</sup>Research Institute of the McGill University Health Centre, Montreal, Quebec. <sup>2</sup>Rotman Research Institute, Baycrest, Toronto, Ontario. <sup>3</sup>Department of Psychology, University of Toronto, Ontario, Canada.

\*Address correspondence to: Bradley R. Buchsbaum, PhD, Rotman Research Institute, Baycrest, 3560 Bathurst Street, Toronto, Ontario M6A 2E1, Canada. E-mail: [bbuchsbaum@research.baycrest.org](mailto:bbuchsbaum@research.baycrest.org)

Decision Editor: Nicole Anderson PhD, CPsych

Received: July 20, 2018; Editorial Decision Date: April 17, 2019

### Abstract

**Objectives:** Aging can reduce the specificity with which memory episodes are represented as distributed patterns of brain activity. It remains unclear, however, whether repeated encoding and retrieval of stimuli modulate this decline. Memory repetition is thought to promote *semanticization*, a transformative process during which episodic memory becomes gradually decontextualized and abstracted. Because semantic memory is considered more resilient to aging than context-rich episodic memory, we hypothesized that repeated retrieval would affect cortical reinstatement differently in young versus older adults.

**Methods:** We reanalyzed data from young and older adults undergoing functional magnetic resonance imaging while repeatedly viewing and recalling short videos. We derived trial-unique multivariate measures of similarity between video-specific brain activity patterns elicited at perception and at recall, which we compared between age groups at each repetition.

**Results:** With repetition, memory representation became gradually more distinct from perception in young adults, as reinstatement specificity converged downward toward levels observed in the older group. In older adults, alternative representations that were item-specific but orthogonal to patterns elicited at perception became more salient with repetition.

**Discussion:** Repetition transformed dominant patterns of memory representation *away* and *orthogonally from* perception in young and older adults, respectively. Although distinct, both changes are consistent with repetition-induced semanticization.

**Keywords:** Aging, Episodic memory, fMRI, MVPA, Semantic memory

The detrimental effects of normal aging on episodic and associative memory are a well-documented phenomenon. However, the advent of modern functional neuroimaging tools offers new means to map these cognitive changes onto dynamic signals that reflect how memory is represented and processed in the brain. In particular, Multivoxel pattern analysis (MVPA) classification techniques (Haxby, 2012; Haynes & Rees, 2006; Mahmoudi, Takerkart, Regragui, Boussaoud, & Brovelli, 2012; Norman, Polyn, Detre, & Haxby, 2006; Tong & Pratte, 2012) applied to functional magnetic

resonance imaging (fMRI) can measure to what extent distributed patterns of brain activity elicited when an event is experienced are “reactivated” when the event is later retrieved from memory. The idea that memory consists of a kind of recapitulation of a state of activation established during perceptual encoding is not new (Damasio, 1989; Descartes, 1892; Gage & Hickok, 2005; Hebb, 2005). The precision with which cortical reinstatement can now be quantified, however, is expanding our understanding of the manner in which the brain represents memory.

In young and healthy populations, MVPA classifiers can capture neural signal elicited at retrieval that distinguishes between object categories (e.g., scenes vs faces or objects; Kuhl, Rissman, Chun, & Wagner, 2011; Polyn, Natu, Cohen, & Norman, 2005), modalities (Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012; Lewis-Peacock & Postle, 2012), encoding conditions (Johnson, McDuff, Rugg, & Norman, 2009; McDuff, Frankel, & Norman, 2009), and even individual stimulus items (Bonnici et al., 2012; Buchsbaum, Lemire-Rodger, Fang, & Abdi, 2012; Chadwick, Hassabis, Weiskopf, & Maguire, 2010; Johnson & Johnson, 2014; Ritchey, Wing, Labar, & Cabeza, 2013; Xue, 2018), providing evidence of reinstatement. Evidence linking the specificity of the reactivated neural pattern to behavioral measures of memory acuity (e.g., recall and recognition performance; McDuff et al., 2009; Ritchey et al., 2013; Staresina, Henson, Kriegeskorte, & Alink, 2012) and to subjective ratings of memory salience (e.g., vividness; Dijkstra, Bosch, & van Gerven, 2017; St-Laurent, Abdi, & Buchsbaum, 2015) have established that neural reinstatement can serve as an index of the sharpness of a memory's neural representation.

A few brain imaging studies that have contrasted memory signal between young and older adults, including our own (St-Laurent, Abdi, Bondad, & Buchsbaum, 2014), have reported an aging-related decrease in the specificity of the memory representation (Abdulrahman, Fletcher, Bullmore, & Morcom, 2017; Johnson, Kuhl, Mitchell, Ankudowich, & Durbin, 2015; McDonough, Cervantes, Gray, & Gallo, 2014; Trelle, Henson, & Simons, Preprint; but see Wang, Johnson, de Chastelaine, Donley, & Rugg, 2016). Such decrease is consistent with a phenomenon known as "de-differentiation" (Burianová, Lee, Grady, & Moscovitch, 2013; Carp, Park, Hebrank, Park, & Polk, 2011; Carp, Park, Polk, & Park, 2011; Grady, 2012; Park et al., 2012; St-Laurent, Abdi, Burianova, & Grady, 2011), whereby older adults show a loss of selectivity or tuning within cortical areas that generally show a degree of functional specialization for sensory, motor, or perceptual features. Of note, the presence of aging-related dedifferentiation at encoding has been shown to predict poor subsequent recall of associative stimuli (Saverino et al., 2016; Zheng et al., 2018). In our study (St-Laurent et al., 2014), we observed a slight reduction in neural pattern specificity during encoding in older compared to young adults (although see Trelle et al., Preprint; Zheng et al., 2018), and a much larger aging-related loss of specificity at retrieval. Although the influence of aging-related encoding factors on subsequent recollection cannot be denied ( Craik & Rose, 2012), the age difference we observed at encoding was too subtle to account entirely for the magnitude of the specificity loss we observed at retrieval.

The aging-related decrease in cortical reinstatement specificity we observed is also consistent with the well-documented reduction in episodic and associative memory performance that is observed in old age (e.g., Henson et al.,

2016; Meusel, Grady, Ebert, & Anderson, 2017; Nyberg & Pudas, 2019; Spaniol & Grady, 2012). An interesting question that arises, however, is whether repeated encoding and retrieval may influence the representational specificity of neural activity patterns in our participants, and how age affects any such memory transformations. One well-known and important tenet in the modern history of memory research is that semantic memory (i.e., knowledge of decontextualized facts) is more resilient to the effects of aging than context-rich episodic memory (Allen, Sliwinski, & Bowie, 2002; Craik & Jennings, 1992; Grady, 2012; Mitchell, 1989; Nilsson, 2003; Nyberg, Backman, Erngrund, Olofsson, & Nilsson, 1996; Rönnlund, Nyberg, Backman, & Nilsson, 2005; Spaniol, Madden, & Voss, 2006) especially when it is relational or associative (Old & Naveh-Benjamin, 2008; Saverino et al., 2016; Spaniol & Grady, 2012). For example, semantic knowledge can mitigate age differences in episodic and associative memory performance in some contexts (e.g., Mohanty, Naveh-Benjamin, & Ratneswar, 2016). Changes in dynamic neural connectivity patterns have also been associated with an increased reliance on crystallized cognition and general knowledge at the expense of specific details in older adults when they recall personal episodic memories (Spreng et al., 2018).

Of note, factors, such as the passage of time (Sekeres et al., 2016; Sekeres, Winocur, Moscovitch, et al., 2018), rehearsal, and repeated retrieval (Antony, Ferreira, Norman, & Wimber, 2017; Moscovitch, Nadel, Winocur, Gilboa, & Rosenbaum, 2006; Nadel & Moscovitch, 1997), are thought to contribute to the transformation of the episodic memory trace into a more semanticized or decontextualized form. According to Trace Transformation Theory (Sekeres, Winocur, & Moscovitch, 2018; Winocur & Moscovitch, 2011), episodic memories can be transformed and reorganized with time and experience into "semantic" representations that retain gist and map onto more general schematic knowledge, but can also lose specific details, including visuospatial elements that make the memory rich and vivid. Within this framework, repeated retrieval could transform the nature of the memory representation with each additional iteration, leading to the gradual formation of a more gist-like representation that is abstracted away from the originally experience, and whose neural signature is increasingly more distinct from the neural patterns first elicited during encoding.

As a memory's specific context gradually becomes blurred with each additional repetition, it is unclear whether salient aging-related differences in memory specificity are maintained or reduced. For example, if episodic memory is more specific and detailed in young compared to older adults, it might undergo a more dramatic transformation as a consequence of iterative retrieval in younger compared to older individuals. To address this idea, we report results from a reanalysis of our data to determine how repeated encoding and retrieval influences patterns of

cortical reinstatement as a function of age. In our previous study (St-Laurent et al., 2014), we compared the ability to reactivate patterns of activity evoked by multimodal video stimuli in older adults (aged 64–78 years), relative to a group of younger adults (aged 21–32 years). Young and older participants were scanned while they alternated between viewing and recalling a set of 11 short video stimuli. In total, each video was seen and retrieved 21 times. As noted, neural patterns of activity were nearly as reliable between the two age groups during stimulus encoding, while neural reactivation *averaged over repetition for each video* showed a marked decrease in older adults' ability to reactivate video-specific activity patterns during retrieval.

For the purpose of addressing the current question, we assessed pattern classification at each trial to determine how, and to what extent, aging and repeated encoding and recall attempts influenced patterns of reactivation. As is often done in studies examining fMRI measures of memory reactivation (see Xue, 2018, for a review), we defined reactivation operationally as the similarity between activity patterns evoked during *encoding* and *retrieval* by a given stimulus. As our study involved 21 repeated encoding and retrieval trials for each stimulus item, we were also able to quantify the reliability of activity patterns across repetitions *within each condition* (the encoding and retrieval conditions, respectively). By assessing activity patterns within condition, we were able to capture video-specific patterns present across repeated retrieval trials, regardless of whether such patterns resemble the ones elicited at encoding. For example, if memory becomes transformed as a function of learning or repetition, activity may become increasingly dissimilar from encoding-related patterns but nevertheless be stable across retrieval attempts. With this approach, we assessed whether the consistency of retrieval-specific patterns evolved differently over trials as a function of age; and, moreover, whether our previous finding of age-related de-differentiation during memory retrieval is evident for all repetitions—with the very first retrieval attempt being of special interest because it more stringently meets the definition of episodic memory as context-specific, that is, of a memory specific to a time and place (Tulving, 2002). We hypothesized that reactivated memories, which are more detailed and specific in young adults, should undergo a steeper decline in pattern specificity as a function of repetition than those of older adults, which may already have a more semantic or transformed signature to begin with.

## Method

### Participants

Nineteen young adults (aged 20–33 years) and 14 older adults (aged 64–78 years) of either sex (see Table 1) were recruited through the Baycrest participant pool, and tested according to a protocol approved by the Rotman

Research Institute's Research Ethics Board. Of the 19 young adults, the data from 14 were previously reported in St-Laurent et al. (2014), and the full set of 19 were studied in St-Laurent et al. (2015). Analyses of data from the 14 older adults were also previously reported in St-Laurent et al. (2014). For the current analysis we used the larger group of 19 younger adults to increase statistical power. Older participants were screened over the phone for dementia with a modified version of the Telephone Interview for Cognitive Status (Welsh, Breitner, & Magruder-Habib, 1993). We used a score of 30/50 as a cutoff point, although the lowest score we observed in our sample was 32. All participants were in good health and had no history of neurological or psychiatric disorder, high blood pressure, or diabetes. All participants had normal hearing, normal color perception, and normal or corrected-to-normal vision.

### fMRI Task

We used a set of short but complex audiovisual videos to elicit brain-wide patterns of activation (Buchsbaum et al., 2012). Fifteen short video clips were gathered from online sites (Vimeo.com and YouTube.com). Four clips were reserved for practice, and 11 were used for the in-scan task (Figure 1). Videos and their soundtracks were edited and exported to an .avi format using iMovie. Each video was associated with a short descriptive title shown in conjunction with the video. This title also served as a retrieval cue during recall trials.

Participants performed a cued recall task while undergoing fMRI scanning. They viewed and mentally replayed the same set of 11 short videos over seven functional runs of 8 min each. Perception (or "encoding") and recall trials were intermixed throughout the runs according to a pseudorandom order that differed between runs and participants. For a particular video, encoding and recall trials always alternated, with the restriction that they were separated by at least two (maximum eight) intervening trials featuring or cueing a different video. Each video was 3.4 s long (plus 0.2 s of buffer time to load), and was shown and recalled 21 times each (3 times per run). All trials were cued by a short title that matched each video (e.g., "Race Car," "Skateboarding Dog"). During encoding trials, a title was shown in blue-green letters (0.9 s) in the center of the screen, followed by the matching video (3.6 s) and an interstimulus interval (ISI) of 0.25–1 s. During recall trials, a title was shown in red letters (0.9 s), followed by a gray rectangle that covered the same portion of the screen as the videos shown on perception trials (4.85 s). Although the gray rectangle was on the screen, participants were instructed to mentally replay the video from memory as vividly and with as many visual and auditory details as possible. Participants were then given 2.75 s to rate the vividness of their recall on a 1–4 scale (1 = not vivid at all; 2–3 = average; 4 = extremely vivid), followed by an ISI of 0–0.75 s. On the basis of piloting, recall trials were 1.25 s

**Table 1.** Peak Coordinates of Recall/Recall Searchlight Analysis

| Older group: Positive repetition trend             |          |          |       |  |               |        |
|--|----------|----------|-------|--|---------------|--------|
| <i>x</i>   | <i>y</i> | <i>z</i> | Hemi  | Label  | <i>z</i> stat | Area   |
| -42.5  | -30      | 6.5      | left  | Planum temporale                             | 4.825018      | 12,582 |
| -45.5  | -60      | 51.5     | left  | Lateral occipital cortex, superior division  | 4.781150      | 12,582 |
| -54.5  | -45      | 30.5     | left  | Supramarginal gyrus, posterior division      | 3.116875      | 12,582 |
| Younger group: negative repetition trend           |          |          |       |  |               |        |
| <i>x</i>   | <i>y</i> | <i>z</i> | Hemi  | Label  | <i>z</i> stat | Area   |
| -57.5  | 0        | -17.5    | left  | Middle temporal gyrus, anterior division     | 4.553251      | 31,536 |
| -57.5  | -63      | 0.5      | left  | Lateral occipital cortex, inferior division  | 4.225403      | 31,536 |
| -48.5  | -60      | 9.5      | left  | Middle temporal gyrus, temporooccipital part | 4.119994      | 31,536 |
| -36.5  | -39      | -14.5    | left  | Temporal fusiform cortex, posterior division | 4.097682      | 31,536 |
| -27.5  | -48      | -5.5     | left  | Lingual gyrus                                | 4.087710      | 31,536 |
| -18.5  | -39      | -2.5     | left  | Cingulate gyrus, posterior division          | 4.023847      | 31,536 |
| -57.5  | -9       | -5.5     | left  | Superior temporal gyrus, anterior division   | 3.882009      | 31,536 |
| -54.5  | -21      | 6.5      | left  | Planum temporale                             | 3.674305      | 31,536 |
| -51.5  | -36      | -5.5     | left  | Middle temporal gyrus, posterior division    | 3.430697      | 31,536 |
| Young > old contrast of linear trend of repetition |          |          |       |  |               |        |
| <i>x</i>   | <i>y</i> | <i>z</i> | Hemi  | Label  | <i>z</i> stat | Area   |
| -42.5  | -60      | 48.5     | left  | Lateral occipital cortex, superior division  | 5.139753      | 83,565 |
| -51.5  | -21      | 6.5      | left  | Heschl's gyrus (includes H1 and H2)          | 4.904393      | 83,565 |
| -39.5  | -54      | 33.5     | left  | Angular gyrus                                | 4.848567      | 83,565 |
| -33.5  | -21      | 9.5      | left  | Insular cortex                               | 4.765009      | 83,565 |
| -51.5  | -36      | 9.5      | left  | Planum temporale                             | 4.658340      | 83,565 |
| -36.5  | -39      | -17.5    | left  | Temporal fusiform cortex, posterior division | 4.102604      | 83,565 |
| -60.5  | -60      | -5.5     | left  | Middle temporal gyrus, temporooccipital part | 3.854847      | 83,565 |
| -9.5   | -60      | 54.5     | left  | Precuneous cortex                            | 3.780770      | 83,565 |
| -51.5  | -69      | 6.5      | left  | Lateral occipital cortex, inferior division  | 3.384498      | 83,565 |
| 41.5   | -18      | 42.5     | right | Postcentral gyrus                            | 4.294886      | 28,107 |
| 23.5   | -6       | -14.5    | right | Right amygdala                               | 3.917757      | 28,107 |
| 62.5   | -30      | -2.5     | right | Middle temporal gyrus, posterior division    | 3.368567      | 28,107 |
| 56.5   | -12      | -29.5    | right | Inferior temporal gyrus, posterior division  | 3.243302      | 28,107 |
| 59.5   | 6        | -11.5    | right | Temporal pole                                | 3.182020      | 28,107 |
| 53.5   | -30      | 18.5     | right | Parietal operculum cortex                    | 3.179239      | 28,107 |
| 35.5   | -48      | 66.5     | right | Superior parietal lobule                     | 4.443508      | 5,832  |
| -24.5  | -9       | 60.5     | left  | Precentral gyrus                             | 4.345305      | 1,323  |
| -18.5  | -39      | -5.5     | left  | Parahippocampal gyrus, posterior division    | 4.143515      | 594    |
| 17.5   | -93      | 24.5     | right | Occipital pole                               | 3.773453      | 54     |
| 26.5   | -66      | -20.5    | right | Occipital fusiform gyrus                     | 3.687650      | 54     |
| 65.5   | -42      | 30.5     | right | Supramarginal gyrus, posterior division      | 3.411494      | 27     |

longer (4.85 s compared to 3.6 s) than the actual duration of the videos to give sufficient time for participants to retrieve the episode from long-term memory. Note that the short ISIs between trials made it very unlikely that subjects would have time to rehearse the kind complex multimodal stimuli used in this study. Before scanning, participants were given instructions and practiced the task using a separate set of four videos. Another practice session with these same four practice videos was completed during an anatomical scan acquired immediately before the first functional run. The

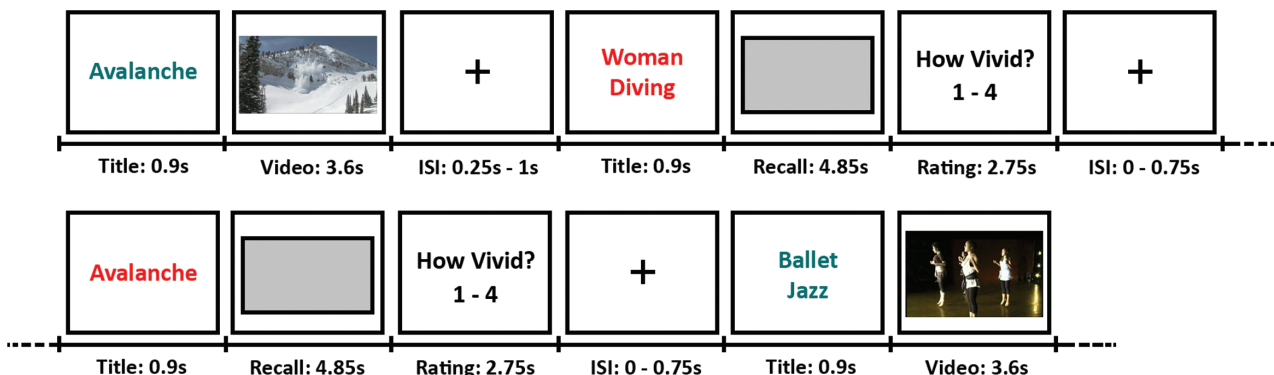
experimental stimulus set was only shown to participants when they performed the in-scan task.

### MRI Scanning and Preprocessing

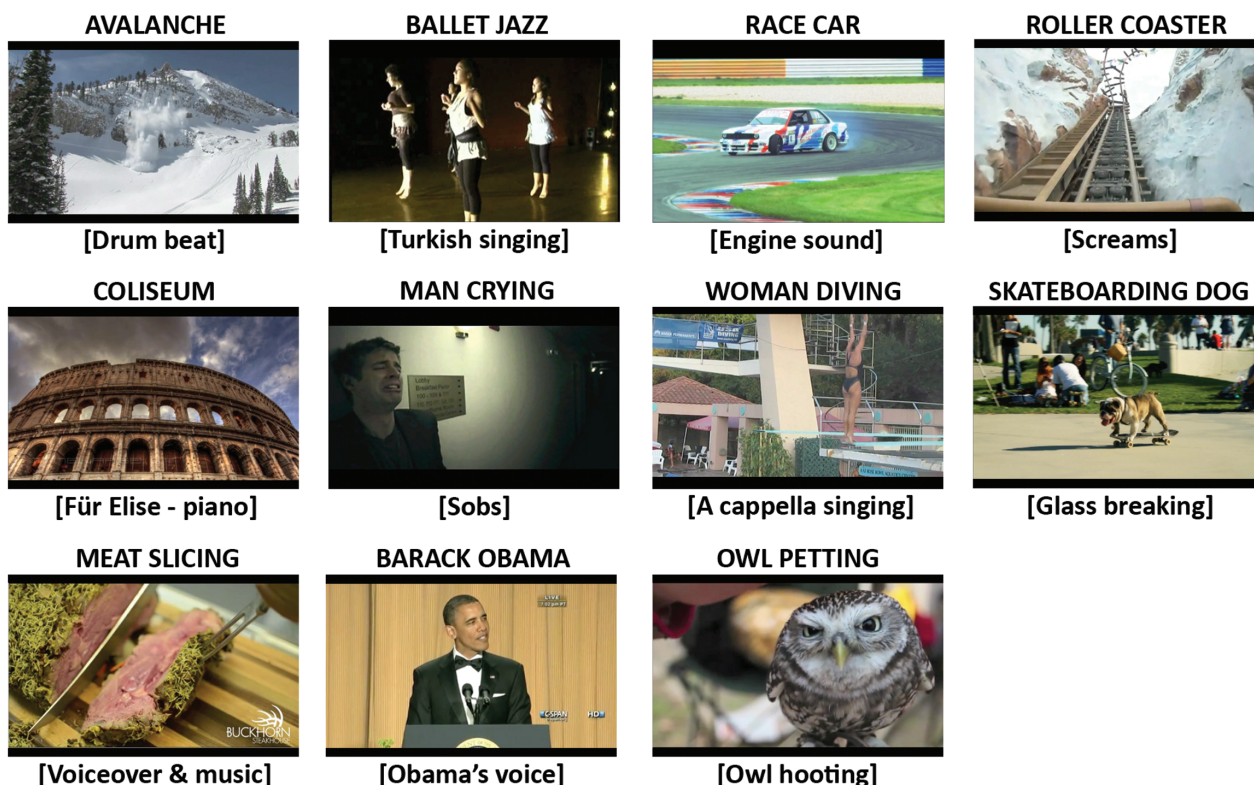
Participants were scanned with a 3.0-T Siemens MAGNETOM Trio MRI scanner using a 12-channel head coil system. High-resolution gradient-echo multi-slice T1-weighted scans (160 slices of 1-mm thickness, 19.2 × 25.6 cm field of view) coplanar with the echo-planar



## In-Scan Task



## Stimulus Set



**Figure 1** Schematic overview of the experimental functional magnetic resonance imaging (fMRI) task from St-Laurent et al. 2014. Top, in-scan cued recall paradigm. Teal-colored titles were always followed by the matching video, whereas red titles were followed by a gray rectangle. Participants mentally replayed the video corresponding to the title while the rectangle was displayed on the screen, and then rated the quality of their recall from 1 to 4. Bottom, title, screenshot, and sound track description for the set of 11 videos viewed and recalled by the participants.

imaging scans (EPIs), as well as whole-brain magnetization prepared rapid gradient-echo (MP-RAGE) 3D T1-weighted scans were first acquired for anatomical localization, followed by T2\*-weighted EPIs sensitive to blood oxygenation level-dependent contrast. EPI images were acquired using a two-shot gradient-echo EPI sequence (22.5 × 22.5 cm field of view with a 96 × 96 matrix size, resulting in an in-plane resolution of 2.35 × 2.35 mm for each of 26 3.5-mm axial slices with a 0.5-mm interslice gap; repetition

time 1.5 s; echo time 27 ms; flip angle 62 degrees). All EPI images were realigned to the average image volume of the first scanning run using *3dvolreg*, a 6-parameter rigid-body algorithm implemented in the AFNI program (Cox, 1996). No spatial smoothing was applied to the EPI images.

Each subject's high-resolution structural MRI was warped to a standard 1 mm MNI template ("MNI152\_T1\_1mm\_brain.nii") shipped with the FSL 5.0 software. Symmetric diffeomorphic nonlinear registration from the

advanced normalization tools (Avants, Epstein, Grossman, & Gee, 2008) was used to register each subject's MRI to the standard template. This transform was then used to warp images in the subject's native space to standardized template space. Note that only the searchlight analysis (described later) made use of these transforms, as the whole-brain classification analyses were conducted on classifier outputs computed in native EPI space.

Multivariate classification analysis generally requires a one-to-one relationship between a data point, or an observation, and a "label." Here, the labels to be classified were the identity of the video encoded or recalled at any particular trial. Because the hemodynamic response to a stimulus event evolves slowly over the course of approximately 10–15 s, several seconds of consecutive *f*MRI data points map onto the label associated with each trial. To estimate the *f*MRI response to each event, we used a multiple regression model that included a hemodynamic regressor for each trial in the experiment ( $21 \times 11$  encoding trials +  $21 \times 11$  retrieval trials = 462 beta estimates per brain voxel). Each trial-wise regressor was generated by convolving a delta function (duration = 3.4 s) with the SPM canonical hemodynamic response function. Owing to the large number of parameters in this model, estimates of the beta coefficients were computed using a form of ridge regression (Mumford, Turner, Ashby, & Poldrack, 2012) as implemented in the AFNI program *3dLSS*. Five additional nuisance parameters were also included in the model to account for low-frequency drift in the *f*MRI signal. Regression models were estimated separately for each run, to avoid any cross-run dependencies in beta estimates. Each subject's 7 runs of 66 beta estimates (total of 462 betas per voxel) were concatenated and the resultant 462 beta estimates per voxel were then used as inputs into multivariate pattern classifiers trained to discriminate distributed activity patterns that could reliably predict which video was viewed or recalled at individual trials.

To define the voxels to be included in the classification analysis, we used a gray matter probability image computed with the FSL program FAST (Smith et al., 2004) using the subjects high-resolution MP-RAGE structural image. The estimated gray matter probability image was downsampled to match the resolution of the EPI data, and thresholded at .33 to create a whole-brain gray matter mask (i.e., a mask included all voxels with a gray matter probability estimate greater than .33).

### Whole-Brain Multivoxel Pattern Analysis

For all MVPA analyses, we used shrinkage discriminant analysis (SDA; Ahdesmäki & Strimmer, 2010), a form of linear discriminant analysis that can be applied to high-dimensional data where there are more variables (i.e., voxels) than observations (i.e., experimental trials). All SDA analyses were performed separately for each subject, and the input data to the classifier were the trial-wise

beta coefficients estimated using the voxel-wise regression method described above.

We trained two SDA pattern classifiers to discriminate between video-specific activation patterns present during encoding and retrieval trials, respectively. Classifier performance was assessed via 7-fold cross-validation (1-fold per *f*MRI run): for each fold, the training set consisted of all the encoding or all the recall trials from 6 runs of data, and the test set consisted of the remaining trials from the held out run. As there were three repetitions of each video per run, class predictions for repetitions 1–3 were derived from a classifier trained on trials from runs 2–7; predictions for repetitions 4–6 were derived from a classifier trained on runs 1 and 3–7; predictions for repetitions 7–9 were derived from a classifier trained on runs 1–2 and 4–7; and so on.

With this cross-validation scheme, we produced *three* different metrics of neural pattern reliability, based on the experimental condition of the trials included in the training and testing sets. We used a classifier trained on encoding trials and tested on the held out encoding trials (Encoding/Encoding) to quantify the consistency of video-specific neural patterns during viewing. We used a classifier trained on recall trials and tested on the held out recall trials (Recall/Recall) to quantify the consistency of video-specific neural patterns during recall, regardless of whether these patterns were modeled on perception. Finally, we used a classifier trained on encoding trials and tested on recall trials from the held out run (Encoding/Recall) to quantify "reactivation," a subjects' capacity to reinstate video-specific activity patterns present at perception while mentally replaying videos from memory.

Each time a trained SDA classifier is applied to a new observation (i.e., a held out experimental trial), it outputs a vector of 11 probabilities with a single value for each label (i.e., one label for each video) in the stimulus set. Each element of the vector is the predicted probability for a different video, with the elements of this probability vector summing to 1. If the highest probability in this vector corresponds to the true label in the test set, then the classifier has correctly predicted which video the subject was viewing or recalling.

Although classification performance can be understood in simple binary terms (correct/incorrect), a probabilistic classifier output that estimates the likelihood of each label provides a continuous—and richer—measure of performance. Even when the classifier is nominally "wrong" (i.e., when the maximal probability does not correspond to the true label), there is still information to be derived by reading out the classifier's estimated probability for the true test label.

### Baseline Classification Model

Because our experiment had 11 videos that that were repeated an equal number of times in each run (3 encoding, 3 recall) a classifier trained on 6 runs to predict the videos

in a held out run has a theoretical chance performance of  $1/11$  (.0909). However, biases in our randomization, systematic head motion, or other unknown dependencies in the data set could cause the chance classification rate to empirically diverge from this theoretical level. To test for this possibility, we replicated the analyses described in the previous section, except instead of using voxels derived from a gray matter mask we created a mask that included only voxels with a cerebrospinal fluid (CSF) probability estimate greater than .95. A model trained only using voxels in the CSF should also be corrupted by the kind of biases associated with head motion, failures of randomization, or other temporal dependencies that would affect the chance accuracy rate in a similar way as in the gray matter analysis. We, therefore, use the CSF classification analysis to confirm that the results of the gray matter classification analysis are not only above the “theoretical chance” level of .0909 but also exceed the chance level calculated using the cross-validated performance metrics derived from the CSF classifier predictions.

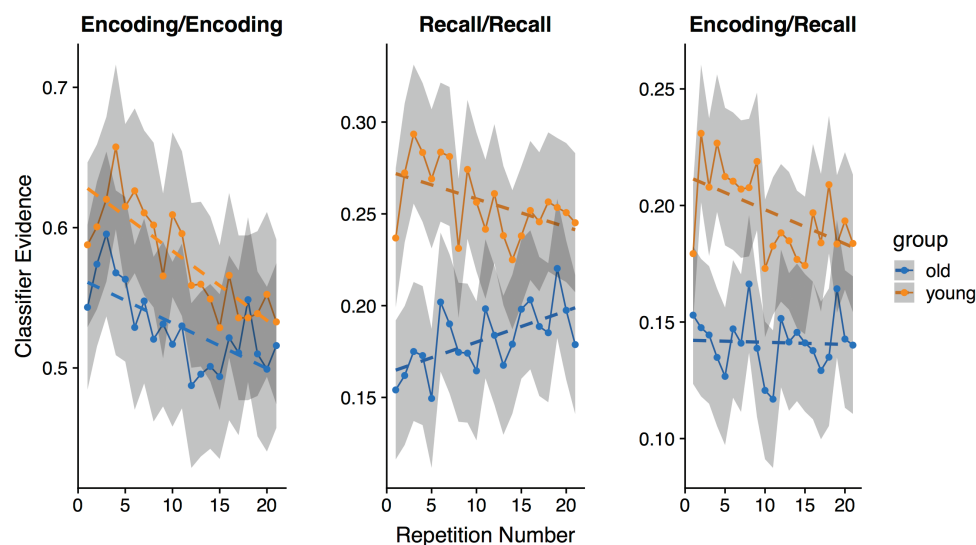
### Linear Mixed-Effect Modeling of Whole-Brain Classifier Outputs

To analyze how neural patterns of activity change with repeated encoding and retrieval attempts, we computed general linear mixed-effect models based on probabilistic classifier output. We derived a separate model for each of our three classification metrics: Encoding/Encoding, Recall/Recall, and Encoding/Recall. In each model, we included a within-subjects fixed effect for the linear trend of repetition (1–21), a between-subjects factor for age group (old/young), and a repetition by age group interaction term. Random effects were also included for the subject (random

intercepts) and repetition (random slope) terms. Because the dependent variable was a probability, we used a logistic regression model, as implemented in the *glmer* function of the *lme4* package in R (Bates, Maechler, Bolker, & Walker, 2015). Each mixed model examined whether video-specific activity patterns changed over repetition and/or group. In each case, a main effect of group indicated an overall difference in the classifier’s ability to predict the correct video between young and older adults; a significant repetition trend indicated that the classifier’s prediction accuracy changed as a function of repetition; finally, a group by repetition interaction indicated that the slope of the accuracy function (as a function of repetition) differed across age groups.

### Searchlight Pattern Analysis

A multivariate searchlight analysis (Kriegeskorte, Goebel, & Bandettini, 2006) was carried out to identify the regions of the brain that underlie the age by repetition crossover interaction trends identified with the Recall/Recall pattern classification conducted on whole brain (see Figure 2). To gain information about the spatial distribution and temporal trends associated with the Recall/Recall classifier, we conducted a multivariate searchlight analysis with an 8-mm spherical radius. The searchlight analysis was conducted in a manner analogous to the whole-brain classification analysis, using the same cross-validation procedure, and producing the same probabilistic classifier outputs. The principal difference between the searchlight analysis and the aforementioned whole-brain analysis was the number of voxels entered in each classifier: the searchlight was repeated for each 8-mm spherical neighborhood around each voxel in the subject’s native EPI space to localize voxels whose activity contributed to classification.



**Figure 2.** Classifier evidence as a function of repetition. Probabilistic classifier evidence for the correct video for training and test set combinations: Encoding/Encoding, Recall/Recall, and Encoding/Recall (reactivation). For reference, chance probability in all three cases is  $1/11 = .0909$ . Shaded confidence bands represent 95% bootstrap confidence intervals.

To examine which regions of the brain are primarily driving the repetition trends observed in the whole-brain analysis of the Recall/Recall classifier (see *Results*), a similar strategy to the one used for the whole-brain analysis was adopted (see *Linear Mixed-Effect Modeling of Whole-Brain Classifier Outputs*). In this case, however, the trend analysis was separately computed for each voxel and could, therefore, be used to isolate regions in the brain where Recall/Recall classifier performance increased or decreased as a function of repetition. In addition, because it was not computationally practical to estimate a linear mixed-effects model for each voxel, we instead estimated trend effects using a two-stage approach. In the first stage, a standard logistic regression model using R's *glm* function was carried out for every voxel in each subject's native EPI space to estimate the effect of repetition on classifier performance, which was again expressed as the classifier's maximum predicted probability for every trial. In each logistic regression model (one for every subject and voxel) the dependent variable was classifier probability for the correct category, and the independent variable was repetition (1–21). The estimated normalized effect of repetition at each voxel was the  $z$ -score derived from the logistic regression model output (i.e., the ratio of the estimated effect normalized by the standard error). The set of voxel-wise  $z$ -scores, therefore, formed, for each subject, a map of estimated trends of the accuracy of the Recall/Recall pattern classifier. These maps were then warped to MNI space and analyzed at the group level using standard voxel-wise  $t$  tests. The resulting statistical maps were then tested for statistical significance using a randomization test implanted in the FSL 5.0 program *randomise*, which implement a stringent correction for multiple comparisons (Eklund, Nichols, & Knutsson, 2016).

## Results

### Age Differences in Classifier Accuracy

Classifier results for the three train-test configurations (Encoding/Encoding, Recall/Recall, Encoding/Recall) are plotted in [Figure 2](#). In all three cases, classification was generally more accurate for young than for older adults. However, only for Recall/Recall and Encoding/Recall classification was the main effect of age group statistically significant: (Encoding/Encoding, main effect of group:  $z = 1.38$ ,  $p = .16$ ; Recall/Recall:  $z = 10.63$ ,  $p < .0001$ ; Encoding/Recall:  $z = 6.1$ ,  $p < .0001$ ). This result is consistent with our previous findings from [St-Laurent et al. \(2014\)](#) where we showed much larger age differences for pattern reactivation during recall than for video-specific activity pattern present at encoding.

### Changes in Whole-Brain Gray Matter Classifier Accuracy Over Repetition

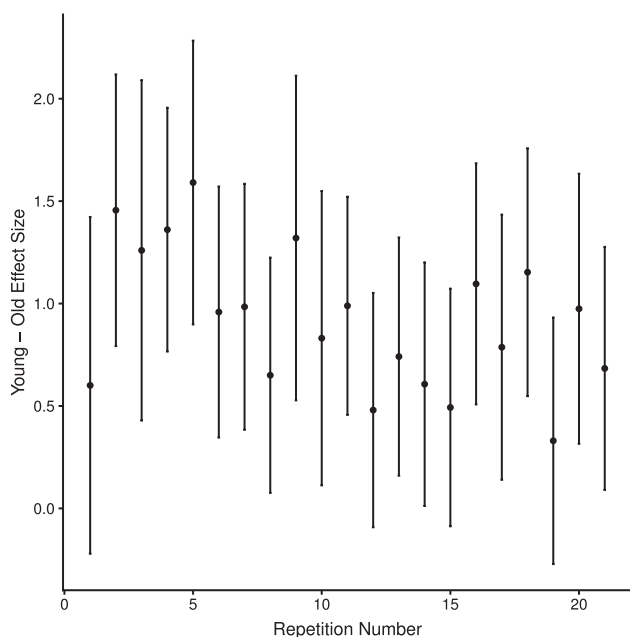
For the Encoding/Encoding model, there was a main effect of repetition such that classifier accuracy declined as a

function of repetition ( $z = -3.8$ ,  $p = .00014$ ). However, there was no group by repetition interaction ( $p = .304$ ), indicating that the rate and magnitude of change was not different for the two age groups. By contrast, there was a significant group by repetition interaction for the Encoding/Recall model ( $z = 2.222$ ,  $p = .0262$ ), indicating a differential slope in classifier accuracy as a function of repetition between age groups. In this case, the negative slope for the young group was significantly different from zero ( $z = -2.838$ ,  $p = .0045$ ), whereas the slope for the old group was not ( $z = .718$ ;  $p = .476$ ). This result indicates that distributed patterns of memory representation became increasingly distinct from patterns elicited at perception in young adults, converging toward levels observed in older adults, which were lower and constant. Finally, the Recall/Recall model also revealed a clear group by repetition interaction ( $z = 3.345$ ,  $p = .0008$ ): classification accuracy decreased as a function of repetition in young adults, whereas it *increased* with repetition in older adults. Although in opposite directions, the simple effect for the linear trend of repetitions was significant within each age group (young:  $z = -2.175$ ,  $p = .02$ ; older:  $z = 2.588$ ,  $p = .0096$ ). Importantly, the Recall/Recall classifier captured the most consistent patterns of stimulus-specific representation elicited at retrieval, regardless of whether such patterns were modeled on perception. With repetition, the salience of these patterns increased in older adults, but decreased in young adults. Thus, whereas classifier accuracy declined as a function of repetition for all three models in the young group, classifier accuracy for the older group declined for the Encoding/Encoding model, increased for the Recall/Recall model, and remained stable for the Encoding/Recall model.

### Effect of Age on Reactivation on the First Trial

A key contribution of the current analyses was the ability to test whether repeated stimulus viewing and stimulus retrieval was responsible for our prior finding ([St-Laurent et al. 2014](#)) in neural pattern reactivation. The aforementioned analyses do not support this interpretation: pattern reactivation (Encoding/Recall classification) declined as a function of repetition in young adults, while it remained constant in older adults. If anything the age effect was attenuated by the repetition component of the task, as it brought young adults' performance closer to levels observed in the older group. To quantify this effect more precisely, we computed 1,000 bootstrap estimates of the effect size (Cohen's  $d$ ) of the young–older difference in Encoding/Recall classifier performance for each of the 21 recall trials. Using the bootstrap distributions, we also computed the 95% confidence intervals for each of these difference scores, yielding bootstrapped estimates of the difference scores and corresponding confidence intervals for each of the 21 repetitions. As is evident in [Figure 3](#), there is a modest decreasing trend in age differences across repetitions. It is also clear that the confidence intervals overlap between the first repetition and





**Figure 3.** Age differences in reactivation over all repetitions. Effect size and 95% confidence intervals for young–old difference in reactivation (Encoding/Recall) as a function of recall repetition number.

subsequent repetitions, offering little statistical evidence supporting the idea that age differences in reactivation are driven by a learning effect that benefits young adults disproportionately, or that repetition distorts age differences in neural reactivation. Nonetheless, it is true that the effect size of the very first repetition (Cohen's  $d = .601$ ) is lower than the effect size for the second repetition (Cohen's  $d = 1.46$ ), although this difference is not statistically significant—as assessed via a general linear mixed-effects model testing for an age group by repetition interaction, and including only the first two repetitions in the experiment: ( $z = .894$ ,  $p = .372$ ). Note, however, that this is a low powered statistical test that is based on the first two repetitions, which consists of only 22 trials per subject (11 per video per repetition). Thus, three conclusions are probably warranted: (a) there is little evidence supporting the contention that age differences in reactivation are driven or exaggerated by learning, and (b) there is suggestive qualitative evidence indicating that young adults show increased pattern reactivation early in the learning phase (between the first and second trial) relative to older adults, while the age differences decrease with additional repetitions (trials 11–21).

### Regional Searchlight Analysis of Age-Related Differences in Recall/Recall Classifier Accuracy

Linear modeling of the change in classification accuracy as a function of repetition revealed a crossover interaction for the Recall/Recall pattern classifier, whereby younger adults showed a significantly negative slope and older adults showed a significant positive slope (see Figure 2, middle

panel). Because the pattern classifier that produced these effects was trained on all gray matter voxels in the brain, we could not identify the particular brain regions driving the interaction. We therefore used a roving multivariate searchlight approach (Kriegeskorte et al., 2006) to train pattern classifiers over 8-mm spherical “patches” spanning the entire brain volume to gain a better understanding of the spatial distribution of repetition-related changes in pattern activity associated with mental recall.

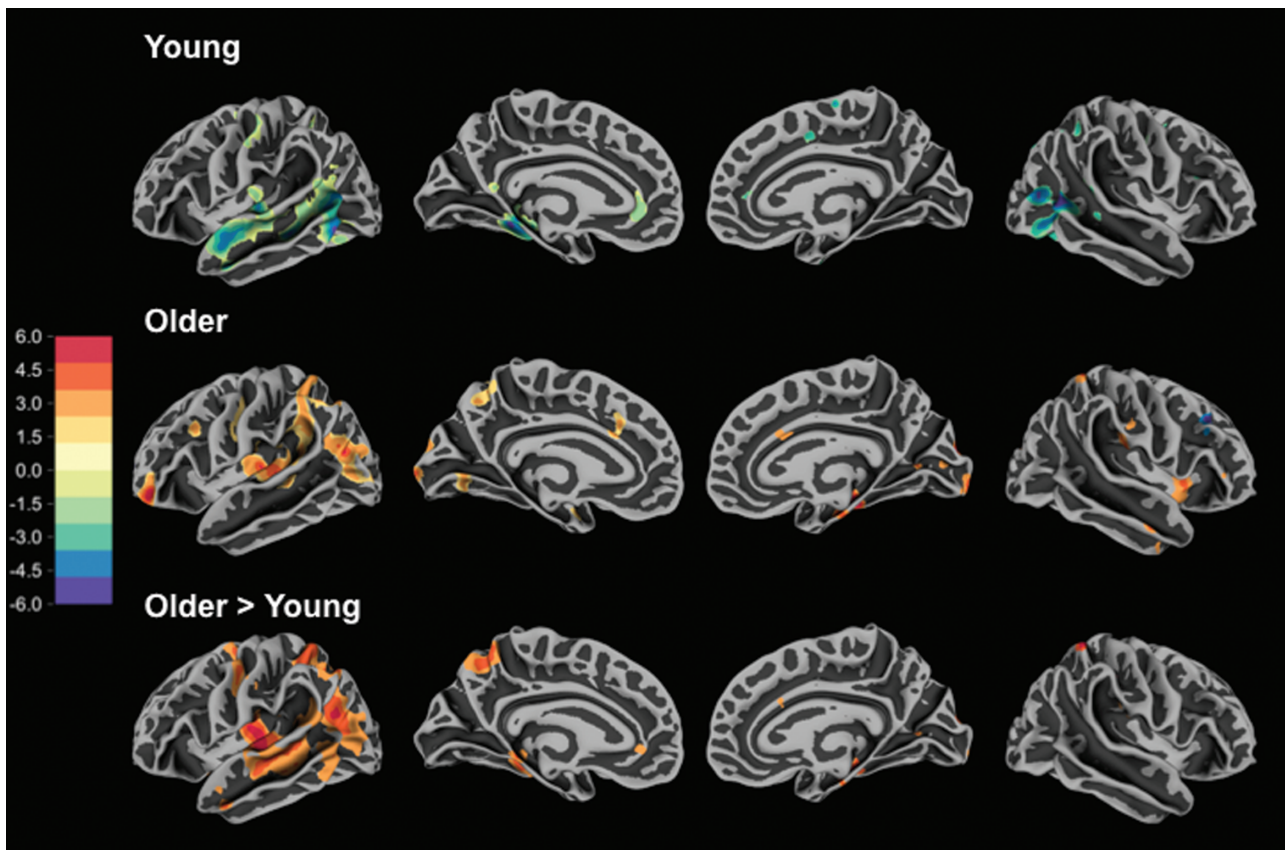
The result of the searchlight analysis revealed that, consistent with the whole-brain classification analysis, younger adults showed only *decreases* in classification accuracy as a function of repetition, whereas older adults primarily showed *increases* in classification accuracy with repetition. As can be seen in Figure 4, these repetition-related decreases in younger adults were concentrated in the left lateral temporal lobe, including the anterior portion of the superior temporal gyrus, the posterior superior temporal sulcus, and the posterior middle temporal gyrus. Right hemisphere decreases were prominent in the posterior lateral occipital region. The repetition-related increases seen in older adults were observed in the posterior temporoparietal regions—including the planum temporale, the supramarginal gyrus, the superior temporal sulcus, and angular gyrus. Increases were also seen in the left precentral gyrus, left ventral anterior prefrontal cortex, and the right hippocampus (see Table 1). A  $t$  test showing regions where the trend differs by group shows a similar pattern (Figure 4, bottom panel; Table 1), although this contrast identifies some regions where younger adults decreased and older adults were stable over repetitions (e.g., the mid portion of the left superior temporal sulcus). Broadly speaking, the major repetition-related effects were observed in regions associated with verbal-semantic processing and language (Binder, Desai, Graves, & Conant, 2009; Hickok & Poeppel, 2007; Price, 2012; Rose, Spalek, & Abdel Rahman, 2015), with a clear concentration in the left hemisphere.

### Comparison of Baseline CSF Classifier to Whole-Brain Gray Matter Classifier

We confirmed that each of the whole-brain gray matter classification models were better able to classify the videos than corresponding models trained only on voxels in the CSF. Averaging over repetitions, we found that the average estimated probability for correct label for the CSF-only trained model ranged between .093 and .107, whereas theoretical chance probability is 1/11 or .0909. A series of two-sample  $t$  tests comparing age groups showed that there was no difference between classification accuracy between the young and the older group for any of the three classifiers (Encoding/Encoding, Encoding/Recall, Recall/Recall; all  $ps > .508$ ).

For the Encoding/Encoding and Recall/Recall classifiers, the average classification accuracies (.107 and .098) were





**Figure 4.** Searchlight analysis of repetition trend for Recall/Recall classifier. Top panel shows regions where Recall/Recall pattern classification accuracy linearly decreased as a function of repetition in the young group. Middle panel shows regions where Recall/Recall pattern classification accuracy linearly increased in older adults. Bottom panel shows regions where the linear trend in pattern classification accuracy was significantly greater for older than for young adults. Colors are in  $t$ -statistic units and are thresholded at  $p < .005$  two-tailed for the purposes of visualization.

significantly greater than the theoretical chance value of .0909 ( $p < .001$ ). However, these biases are small compared to the overall performance of the gray matter classifier, where the average Encoding/Encoding probability was .556 (odds ratio = 10.45 with respect to CSF) and the average Recall/Recall probability was .223 (odds ratio = 2.64 with respect to CSF). Directly comparing the gray matter and classifiers separately with paired  $t$  tests for each group and classifier type yielded 6 separate comparisons (3 classifiers  $\times$  2 age groups). All comparisons showed that, regardless of the group or classifier type, gray matter classifier performance was always significantly superior than predictive performance using the corresponding CSF classifier (all  $ps < 4e-05$ ).

#### Effect of Head Motion and Number of Gray Matter Voxels

We also examined whether two other potential factors were associated with classifier accuracy: namely, subject head motion and the number of gray matter voxels used in the whole-brain classifier. To test for the influence of head motion, we examined whether the maximum frame-wise

displacement, derived from AFNI's *3dvolreg* motion correction program, differed between age groups. Older adults moved slightly more than younger adults, but the difference was not statistically significant ( $t = 1.88, p = .0788$ ). In addition, maximum frame-wise displacement was not a significant predictor of average classification evidence for any of the three models (all  $ps > .161$ ). Nor did motion interact with group in any of the three models (all  $ps > .103$ ). Thus, we conclude that motion had little impact on its own or as a modulatory variable on classification accuracy.

Younger adults had more contributing gray matter voxels entering the whole-brain classifier than older adults: 36,077 versus 29,262, respectively; a difference that is statistically significant ( $p < .001$ ). However, in none of the three models (Encoding/Encoding, Encoding/Recall, Recall/Recall) did we find that the number of gray matter voxels was a significant predictor of classification accuracy after accounting for the main effect of age group (all  $ps > .132$ ). Nor did we find that voxel count interacted with age group in any of the three classification models (all  $ps > .504$ ). In other words, even though the number of gray voxels differed between groups, this difference did not explain any additional variation in classifier performance.

## Discussion

We have reported that older adults were impaired relative to younger adults in reactivating patterns of brain activity during cued mental recall of short audiovisual episodes (St-Laurent et al., 2014). These results suggest that age-related decrease in the quality and precision of episodic memory is associated with corresponding decreases in the specificity of neural reactivation during cued retrieval of complex experiences (see also Abdulrahman et al., 2017; Dennis, Bowman, & Peterson, 2014; Johnson et al., 2015; McDonough et al., 2014; Trelle et al., Preprint; but see Wang et al., 2016). Our previous report, however, was based on an analysis that pooled reactivation measures over a series of 21 repetitions per stimuli. Knowing that repetition can “tamper” with the memory trace (Nadel & Moscovitch, 1997; Sekeres, Winocur, & Moscovitch, 2018; Winocur & Moscovitch, 2011), these results left open the possibility that the age differences we reported reflected an impairment in *learning* to reactivate brain patterns (with young adults learning at a faster rate, and to a greater extent, than older adults with subsequent repetitions), rather than a more standard reactivation deficit observed with a trial-unique protocol. From a methodological standpoint, stimulus repetition provides important advantages when the aim is to capture neural signal that is stimulus-specific: for example, many independent trials are required per item to train a pattern classifier that discriminates among exemplars from the same category. Nevertheless, it is good practice for investigators to assess and report how repetition may have influenced their results, most importantly when studying context-specific episodic memory.

With the current reanalysis, we illustrated to what extent, and in which manner, repeated encoding and retrieval influenced stimulus-specific memory representation in distributed patterns of activity as a function of age. We showed, first, that stimulus specificity was reduced in older adults throughout our task, although this difference decreased somewhat as a function of repetition due to a gradual decline observed in the young group. This pattern is contrary the proposition that different (positive) learning rates over repetitions could account for age differences in cortical reinstatement in our previous results (as hypothesized by Wang et al., 2016). If anything, repetition did not improve the precision of memory representation past the first few trials in our data; instead, it caused a downward change in cortical reinstatement in young adults that pushed the memory representation further away from *true reinstatement* closely modeled on perception, while reinstatement remained constant in older participants. This result is consistent with evidence that retrieval practice enhances memory accessibility, but not quality (Sutterer & Awh, 2016).

This downward trend in reinstatement specificity is consistent with the idea that repeated retrieval contributes to the *semanticization* of the memory trace (Sekeres,

Winocur, & Moscovitch, 2018; Winocur & Moscovitch, 2011) that is, that each subsequent recall helps along the process by which context-specific, experience-near memory (and its neural representation) is transformed into decontextualized, generalized information that is integrated into a broader knowledge structure that supports what we understand about the world (Conway, 2009; Conway & Pleydell-Pearce, 2000), that is, what we call semantic memory (Tulving, 1985, 2002). As noted, we observed age differences in memory representation from the very first repetition onward, which is consistent with the well-documented fact that episodic and associative memory declines with old age. What emerged next, over each subsequent repetition, is telling: performance by the Encoding/Recall classifier, which quantifies how closely the reactivated pattern is modeled on perception, declined in the young group, suggesting a more general or abstracted level of representation. In fact, specificity levels converged toward those observed in older adults, in whom neural reinstatement remained low but constant across repetitions.

Interestingly, Recall/Recall classifier performance *increased* over repetitions in the older group, revealing the gradual emergence or sharpening of a consistent form of memory representation that was orthogonal to patterns elicited at encoding (otherwise Encoding/Recall classifier performance would also increase). This representation may reflect the strengthening of a transformed, processed or otherwise semanticized memory representation that is consistent and item-specific, but not a faithful perceptual reactivation per se. A regional searchlight analysis further elucidated the brain regions that contributed most to the repetition effects identified in the Recall/Recall whole-brain classifier. Young adults showed only negative changes in the classifier accuracy as a function of repetition, and these were most evident in superior portion of the left lateral temporal lobe. In contrast, older adults showed increasing pattern classification in regions also in the left hemisphere but generally more dorsally and posteriorly situated. Indeed, it appears the negative-going activation patterns in younger adults and the positive-going activation patterns in older adults are in rough accordance with the auditory-verbal ventral and dorsal streams, respectively (Buchsbaum et al., 2005a; Hickok & Poeppel, 2004; Rauschecker & Scott, 2009). It has been proposed that the auditory dorsal stream is critical for sensorimotor control of language, which is most critical for the production, repetition, and imitation of temporally structured stimuli, including spoken language (Buchsbaum et al., 2011), sign language (Buchsbaum et al., 2005b), and music (Hickok, Buchsbaum, Humphries, & Muftuler, 2003). Thus, one potential explanation for the increase in dorsal stream pattern activity as a function of repetition in older adults is that they are increasingly relying on a verbal-sensorimotor—rather than a more veridical ventral-stream mediated “what” representation—representations of the temporal structure of the videos to maintain and rehearse them in memory.

Another region of the brain that increased pattern classification as a function of repetition in older adults was the left angular gyrus, a region commonly associated with semantic cognition (e.g., Binder et al., 2009). It is generally accepted that general knowledge, crystallized cognition and other forms of semantic knowledge are typically well preserved well into old age (Allen et al., 2002; Mitchell, 1989; Nyberg & Pudas, 2019; Rönnlund et al., 2005; Spaniol et al., 2006), and older adults have been shown to rely on these faculties to compensate for episodic memory deficits in some contexts (Craik & Bialystok, 2006; Mohanty et al., 2016; Spreng et al., 2018). Here, our results indicate that older adults may have relied on their preserved semantic memory to construct and reactivate item-specific representations during recall that were unrelated to perception.

Of note, Recall/Recall classifier performance *decreased* over repetition in our young group, that is, the consistency of the item-specific neural representations they activated at retrieval—whether or not it was modeled on perception—declined over trials. This decline could simply suggest a gradual disengagement with the task (i.e., fatigue), although the absence of a similar effect in the older group makes us question this interpretation. Instead, this result could indicate a switch in representation (from specific to abstract or transformed) that is more drastic than in the older group. The Recall/Recall classifier was optimized to pick up whichever patterns were most “average” or “consistent” over trials. Initially, young adults retrieved highly specific and detailed memory representations (as shown by the Encoding/Recall classifier); a gradual switch toward a more processed or semantic form of mental representation, if too dramatic, may have hindered the performance of a Recall/Recall classifier optimized to pick up patterns that recapitulated perception. If, on the other hand, a more semantic or transformed type of representation was favored from the onset of the task, the Recall/Recall classifier would be “tuned” to pick up this “semanticized” pattern instead, and its performance over repetition would fluctuate according to the sharpness of this pattern at recall, as seen in our older group. Indeed, a significantly closer resemblance between the Encoding/Recall and Recall/Recall searchlight maps averaged over repetition in young compared to older adults, as revealed when correlating the two searchlight volumetric classification maps, supports this interpretation ( $t(27) = 4.1, p = .0003$ ; mean  $r = 0.8634$  for young adults; mean  $r = 0.5028$  for older adults). This analysis strongly suggests that classification accuracy declined over repetitions within similar sets of brain regions for the Encoding/Recall and Recall/Recall classifiers in the young group, supporting the idea that Recall/Recall classifiers were tuned to pick up patterns that were more closely modeled on perception in young adults. Nevertheless, the current paradigm could not dissociate fatigue or other temporal processes from the transformation of the memory trace in our younger group, highlighting the needs for a paradigm that can resolve these potential explanations. Moreover, there is reason to believe that even more dramatic changes occur in memory representations where retrieval is tested hours or days after

the initial encoding (Lee, Kravitz, & Baker, 2018; Sneve et al., 2015; Vidal-Piñeiro et al., 2017).

## Conclusion

Our results demonstrate that reasonably accurate and stable estimates of item-specific reactivation can be obtained with paradigms that include multiple item repetitions with multiple retrieval attempts for those items. Some changes in activity patterns consistent with the emergence of abstracted or semantic forms of memory representation were observed, however, especially in the older group. Older adults showed overall less reactivation than younger adults throughout the experiment, consistent with the well-known age-related decline in the precision of memory. However, the consistency of their recall-related memory patterns increased as a function of repetition—even though these patterns did not resemble encoding-related activation patterns. We interpret this finding as indicating a transformation process, whereby older adults adaptively encode a repeated event into a stable, semanticized, or low-resolution form that discards some of the perceptual detail of the fully encoded stimulus.

Overall, item-specific pattern classification based on multiple trials is a valid approach to understand the interplay of neural events that give rise to episodic memory, while it can also be used to determine the influence of time and experimental manipulations on the transformation of memory trace. In general, however, we recommend that experimenters interested in the methodological advantages offered by stimulus repetition quantify its effects on their neural patterns of interest.

## Funding

Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Award to B. R. Buchsbaum. Canadian Institutes of Health Research (CIHR) Grant 158279 to B. R. Buchsbaum.

## Conflict of Interest

None reported.

## References

- Abdulrahman, H., Fletcher, P. C., Bullmore, E., & Morcom, A. M. (2017). Dopamine and memory dedifferentiation in aging. *Neuroimage*, *153*, 211–220. doi:10.1016/j.neuroimage.2015.03.031
- Ahdesmäki, M., & Strimmer, K. (2010). Feature selection in omics prediction problems using cat scores and false non-discovery rate control. *Annals of Applied Statistics*, *4*, 503–519. doi:10.1214/09-AOAS277
- Allen, P. A., Sliwinski, M., & Bowie, T. (2002). Differential age effects in semantic and episodic memory, part II: Slope and intercept analyses. *Experimental Aging Research*, *28*, 111–142. doi:10.1080/03610730252800157



- Antony, J. W., Ferreira, C. S., Norman, K. A., & Wimber, M. (2017). Retrieval as a fast route to memory consolidation. *Trends in Cognitive Sciences*, *21*, 573–576. doi:10.1016/j.tics.2017.05.001
- Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*, 26–41. doi:10.1016/j.media.2007.06.004
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. doi:10.18637/jss.v067.i01
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex (New York, N.Y.: 1991)*, *19*, 2767–2796. doi:10.1093/cercor/bhp055
- Bonnici, H. M., Chadwick, M. J., Lutti, A., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2012). Detecting representations of recent and remote autobiographical memories in vmPFC and hippocampus. *Journal of Neuroscience*, *32*, 16982–16991. doi:10.1523/JNEUROSCI.2475-12.2012
- Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory—An aggregate analysis of lesion and fMRI data. *Brain and Language*, *119*, 119–128. doi:10.1016/j.bandl.2010.12.001
- Buchsbaum, B. R., Lemire-Rodger, S., Fang, C., & Abdi, H. (2012). The neural basis of vivid memory is patterned on perception. *Journal of Cognitive Neuroscience*, *24*, 1867–1883. doi:10.1162/jocn\_a\_00253
- Buchsbaum, B. R., Olsen, R. K., Koch, P. F., Kohn, P., Kippenhan, J. S., & Berman, K. F. (2005a). Reading, hearing, and the planum temporale. *Neuroimage*, *24*, 444–454. doi:10.1016/j.neuroimage.2004.08.025
- Buchsbaum, B., Pickell, B., Love, T., Hatrak, M., Bellugi, U., & Hickok, G. (2005b). Neural substrates for verbal working memory in deaf signers: fMRI study and lesion case report. *Brain and Language*, *95*, 265–272. doi:10.1016/j.bandl.2005.01.009
- Burianová, H., Lee, Y., Grady, C. L., & Moscovitch, M. (2013). Age-related dedifferentiation and compensatory changes in the functional network underlying face processing. *Neurobiology of Aging*, *34*, 2759–2767. doi:10.1016/j.neurobiolaging.2013.06.016
- Carp, J., Park, J., Hebrank, A., Park, D. C., & Polk, T. A. (2011). Age-related neural dedifferentiation in the motor system. *PLoS One*, *6*, e29411. doi:10.1371/journal.pone.0029411
- Carp, J., Park, J., Polk, T. A., & Park, D. C. (2011). Age differences in neural distinctiveness revealed by multi-voxel pattern analysis. *Neuroimage*, *56*, 736–743. doi:10.1016/j.neuroimage.2010.04.267
- Chadwick, M. J., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2010). Decoding individual episodic memory traces in the human hippocampus. *Current Biology*, *20*, 544–547. doi:10.1016/j.cub.2010.01.053
- Conway, M. A. (2009). Episodic memories. *Neuropsychologia*, *47*, 2305–2313. doi:10.1016/j.neuropsychologia.2009.02.003
- Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*, 261–288. doi:10.1037/0033-295X.107.2.261
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*, 162–173.
- Craik, F. I., & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences*, *10*, 131–138. doi:10.1016/j.tics.2006.01.007
- Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik, & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition* (pp. 51–110). Hillsdale, NJ: Erlbaum.
- Craik, F. I., & Rose, N. S. (2012). Memory encoding and aging: A neurocognitive perspective. *Neuroscience and Biobehavioral Reviews*, *36*, 1729–1739. doi:10.1016/j.neubiorev.2011.11.007
- Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*, 25–62. doi:10.1016/0010-0277(89)90005-X
- Dennis, N. A., Bowman, C. R., & Peterson, K. M. (2014). Age-related differences in the neural correlates mediating false recollection. *Neurobiology of Aging*, *35*, 395–407. doi:10.1016/j.neurobiolaging.2013.08.019
- Descartes, R. (1892). *The Philosophy of descartes: In extracts from his writing*. New York, NY: H. Holt
- Dijkstra, N., Bosch, S. E., & van Gerven, M. A. (2017). Vividness of visual imagery depends on the neural overlap with perception in visual areas. *Journal of Neuroscience*, *37*, 1367–1373. doi:10.1523/JNEUROSCI.3022-16.2016
- Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 7900–7905. doi:10.1073/pnas.1602413113
- Gage, N., & Hickok, G. (2005). Multiregional cell assemblies, temporal binding and the representation of conceptual knowledge in cortex: A modern theory by a “classical” neurologist, Carl Wernicke. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior*, *41*, 823–832. doi:10.1016/S0010-9452(08)70301-0
- Grady, C. (2012). The cognitive neuroscience of ageing. *Nature reviews. Neuroscience*, *13*, 491–505. doi:10.1038/nrn3256
- Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: The early beginnings. *Neuroimage*, *62*, 852–855. doi:10.1016/j.neuroimage.2012.03.016
- Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews. Neuroscience*, *7*, 523–534. doi:10.1038/nrn1931
- Hebb, D. O. (2005). The first stage of perception: Growth of the assembly. In *The Organization of Behavior: A Neuropsychological theory* (pp. 102–120). London, UK: Psychology Press. doi:10.4324/9781410612403
- Henson, R. N., Campbell, K. L., Davis, S. W., Taylor, J. R., Emery, T., Erzincinoglu, S., . . . Kievit, R. A.; Cam-CAN. (2016). Multiple determinants of lifespan memory differences. *Scientific Reports*, *6*, 32527. doi:10.1038/srep32527
- Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, *15*, 673–682. doi:10.1162/089892903322307393

- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*, 67–99. doi:10.1016/j.cognition.2003.10.011
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, *8*, 393–402. doi:10.1038/nrn2113
- Johnson, M. R., & Johnson, M. K. (2014). Decoding individual natural scene representations during perception and imagery. *Frontiers in Human Neuroscience*, *8*, 59. doi:10.3389/fnhum.2014.00059
- Johnson, M. K., Kuhl, B. A., Mitchell, K. J., Ankudowich, E., & Durbin, K. A. (2015). Age-related differences in the neural basis of the subjective vividness of memories: Evidence from multivoxel pattern classification. *Cognitive, Affective & Behavioral Neuroscience*, *15*, 644–661. doi:10.3758/s13415-015-0352-9
- Johnson, J. D., McDuff, S. G., Rugg, M. D., & Norman, K. A. (2009). Recollection, familiarity, and cortical reinstatement: A multivoxel pattern analysis. *Neuron*, *63*, 697–708. doi:10.1016/j.neuron.2009.08.011
- Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 3863–3868. doi:10.1073/pnas.0600244103
- Kuhl, B. A., Rissman, J., Chun, M. M., & Wagner, A. D. (2011). Fidelity of neural reactivation reveals competition between memories. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 5903–5908. doi:10.1073/pnas.1016939108
- Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2018). Differential representations of perceived and retrieved visual information in hippocampus and cortex. *Cerebral Cortex* doi:10.1093/cercor/bhy325
- Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, *24*, 61–79. doi:10.1162/jocn\_a\_00140
- Lewis-Peacock, J. A., & Postle, B. R. (2012). Decoding the internal focus of attention. *Neuropsychologia*, *50*, 470–478. doi:10.1016/j.neuropsychologia.2011.11.006
- Mahmoudi, A., Takerkart, S., Regragui, F., Boussaoud, D., & Brovelli, A. (2012). Multivoxel pattern analysis for fMRI data: A review. *Computational and Mathematical Methods in Medicine*, *2012*, 961257. doi:10.1155/2012/961257
- McDonough, I. M., Cervantes, S. N., Gray, S. J., & Gallo, D. A. (2014). Memory's aging echo: Age-related decline in neural reactivation of perceptual details during recollection. *Neuroimage*, *98*, 346–358. doi:10.1016/j.neuroimage.2014.05.012
- McDuff, S. G., Frankel, H. C., & Norman, K. A. (2009). Multivoxel pattern analysis reveals increased memory targeting and reduced use of retrieved details during single-agenda source monitoring. *Journal of Neuroscience*, *29*, 508–516. doi:10.1523/JNEUROSCI.3587-08.2009
- Meusel, L. A., Grady, C. L., Ebert, P. E., & Anderson, N. D. (2017). Brain-behavior relationships in source memory: Effects of age and memory ability. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior*, *91*, 221–233. doi:10.1016/j.cortex.2016.12.023
- Mitchell, D. B. (1989). How many memory systems? Evidence from aging. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *15*, 31–49. doi:10.1037//0278-7393.15.1.31
- Mohanty, P. P., Naveh-Benjamin, M., & Ratneshwar, S. (2016). Beneficial effects of semantic memory support on older adults' episodic memory: Differential patterns of support of item and associative information. *Psychology and Aging*, *31*, 25–36. doi:10.1037/pag0000059
- Moscovitch, M., Nadel, L., Winocur, G., Gilboa, A., & Rosenbaum, R. S. (2006). The cognitive neuroscience of remote episodic, semantic and spatial memory. *Current Opinion in Neurobiology*, *16*, 179–190. doi:10.1016/S0959-4388(06)00038-9 [pii]
- Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage*, *59*, 2636–2643. doi:10.1016/j.neuroimage.2011.08.076
- Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, *7*, 217–227. doi:10.1016/S0959-4388(97)80010-4
- Nilsson, L. G. (2003). Memory function in normal aging. *Acta Neurologica Scandinavica. Supplementum*, *179*, 7–13. doi:10.1034/j.1600-0404.107.s179.5.x
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*, 424–430. doi:10.1016/j.tics.2006.07.005
- Nyberg, L., Bäckman, L., Erngrund, K., Olofsson, U., & Nilsson, L. G. (1996). Age differences in episodic memory, semantic memory, and priming: Relationships to demographic, intellectual, and biological factors. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *51*, 234–240. doi:10.1093/geronb/51b.4.p234
- Nyberg, L., & Pudas, S. (2019). Successful memory aging. *Annual Review of Psychology*, *70*, 219–243. doi:10.1146/annurev-psych-010418-103052
- Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology Aging*, *23*, 104–118. doi:10.1037/0882-7974.23.1.104
- Park, J., Carp, J., Kennedy, K. M., Rodrigue, K. M., Bischof, G. N., Huang, C. M., . . . Park, D. C. (2012). Neural broadening or neural attenuation? Investigating age-related dedifferentiation in the face network in a large lifespan sample. *Journal of Neuroscience*, *32*, 2154–2158. doi:10.1523/JNEUROSCI.4494-11.2012
- Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science (New York, N.Y.)*, *310*, 1963–1966. doi:10.1126/science.1117645
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, *62*, 816–847. doi:10.1016/j.neuroimage.2012.04.062
- Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, *12*, 718–724. doi:10.1038/nn.2331
- Ritche, M., Wing, E. A., LaBar, K. S., & Cabeza, R. (2013). Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cerebral Cortex (New York, N.Y.: 1991)*, *23*, 2818–2828. doi:10.1093/cercor/bhs258



- Rönnlund, M., Nyberg, L., Bäckman, L., & Nilsson, L. G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology and Aging, 20*, 3–18. doi:10.1037/0882-7974.20.1.3
- Rose, S. B., Spalek, K., & Abdel Rahman, R. (2015). Listening to puns elicits the co-activation of alternative homophone meanings during language production. *PLoS One, 10*, e0130853. doi:10.1371/journal.pone.0130853
- Saverino, C., Fatima, Z., Sarraf, S., Oder, A., Strother, S. C., & Grady, C. L. (2016). The associative memory deficit in aging is related to reduced selectivity of brain activity during encoding. *Journal of Cognitive Neuroscience, 28*, 1331–1344. doi:10.1162/jocn\_a\_00970
- Sekeres, M. J., Bonasia, K., St-Laurent, M., Pishdadian, S., Winocur, G., Grady, C., & Moscovitch, M. (2016). Recovering and preventing loss of detailed memory: Differential rates of forgetting for detail types in episodic memory. *Learning & Memory (Cold Spring Harbor, N.Y.), 23*, 72–82. doi:10.1101/lm.039057.115
- Sekeres, M. J., Winocur, G., & Moscovitch, M. (2018). The hippocampus and related neocortical structures in memory transformation. *Neuroscience Letters, 680*, 39–53. doi:10.1016/j.neulet.2018.05.006
- Sekeres, M. J., Winocur, G., Moscovitch, M., Anderson, J. A. E., Pishdadian, S., Martin Wojtowicz, J., . . . Grady, C. L. (2018). Changes in patterns of neural activity underlie a time-dependent transformation of memory in rats and humans. *Hippocampus, 28*, 745–764. doi:10.1002/hipo.23009
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., . . . Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage, 23*(Suppl 1), S208–219. doi:10.1016/j.neuroimage.2004.07.051
- Sneve, M. H., Grydeland, H., Nyberg, L., Bowles, B., Amlie, I. K., Langnes, E., . . . Fjell, A. M. (2015). Mechanisms underlying encoding of short-lived versus durable episodic memories. *Journal of Neuroscience, 35*, 5202–5212. doi:10.1523/JNEUROSCI.4434-14.2015
- Spaniol, J., & Grady, C. (2012). Aging and the neural correlates of source memory: Over-recruitment and functional reorganization. *Neurobiology of Aging, 33*, 425 e423–418. doi:S0197-4580(10)00428-8
- Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32*, 101–117. doi:10.1037/0278-7393.32.1.101
- Spreng, R. N., Lockrow, A. W., DuPre, E., Setton, R., Spreng, K. A. P., & Turner, G. R. (2018). Semanticized autobiographical memory and the default—Executive coupling hypothesis of aging. *Neuropsychologia, 110*, 37–43. doi:10.1016/j.neuropsychologia.2017.06.009
- Staresina, B. P., Henson, R. N., Kriegeskorte, N., & Alink, A. (2012). Episodic reinstatement in the medial temporal lobe. *Journal of Neuroscience, 32*, 18150–18156. doi:10.1523/JNEUROSCI.4156-12.2012
- St-Laurent, M., Abdi, H., Bondad, A., & Buchsbaum, B. R. (2014). Memory reactivation in healthy aging: Evidence of stimulus-specific dedifferentiation. *Journal of Neuroscience, 34*, 4175–4186. doi:10.1523/JNEUROSCI.3054-13.2014
- St-Laurent, M., Abdi, H., & Buchsbaum, B. R. (2015). Distributed patterns of reactivation predict vividness of recollection. *Journal of Cognitive Neuroscience, 27*, 2000–2018. doi:10.1162/jocn\_a\_00839
- St-Laurent, M., Abdi, H., Burianová, H., & Grady, C. L. (2011). Influence of aging on the neural correlates of autobiographical, episodic, and semantic memory retrieval. *Journal of Cognitive Neuroscience, 23*, 4150–4163. doi:10.1162/jocn\_a\_00079
- Sutterer, D. W., & Awh, E. (2016). Retrieval practice enhances the accessibility but not the quality of memory. *Psychonomic Bulletin & Review, 23*, 831–841. doi:10.3758/s13423-015-0937-x
- Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology, 63*, 483–509. doi:10.1146/annurev-psych-120710-100412
- Trelle, A. N., Henson, R. N., & Simons, J. S. (Preprint). Neural evidence for age differences in representational quality and strategic retrieval processes. *BioRxiv*. doi:10.1101/333484
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26*, 1–12. doi:10.1037/h0080017
- Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*, 1–25. doi:10.1146/annurev.psych.53.100901.135114
- Vidal-Piñeiro, D., Sneve, M. H., Storsve, A. B., Roe, J. M., Walhovd, K. B., & Fjell, A. M. (2017). Neural correlates of durable memories across the adult lifespan: Brain activity at encoding and retrieval. *Neurobiology of Aging, 60*, 20–33. doi:10.1016/j.neurobiolaging.2017.08.017
- Wang, T. H., Johnson, J. D., de Chastelaine, M., Donley, B. E., & Rugg, M. D. (2016). The effects of age on the neural correlates of recollection success, recollection-related cortical reinstatement, and post-retrieval monitoring. *Cerebral Cortex (New York, N.Y.: 1991), 26*, 1698–1714. doi:10.1093/cercor/bhu333
- Welsh, K. A., Breitner, J. C. S., & Magruder-Habib, K. M. (1993). Detection of dementia in the elderly using telephone screening of cognitive status. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 6*, 103–110.
- Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society, 17*, 766–780. doi:10.1017/S1355617711000683
- Xue, G. (2018). The neural representations underlying human episodic memory. *Trends in Cognitive Sciences, 22*, 544–561. doi:10.1016/j.tics.2018.03.004
- Zheng, L., Gao, Z., Xiao, X., Ye, Z., Chen, C., & Xue, G. (2018). Reduced fidelity of neural representation underlies episodic memory decline in normal aging. *Cerebral Cortex (New York, N.Y.: 1991), 28*, 2283–2296. doi:10.1093/cercor/bhx130