



Research report

Neural mechanisms of the spacing effect in episodic memory: A parallel EEG and fMRI study



Xiao Zhao ^{a,b}, Changming Wang ^{a,b}, Qi Liu ^{a,b}, Xiaoqian Xiao ^{a,b},
Ting Jiang ^c, Chuansheng Chen ^d and Gui Xue ^{a,b,*}

^a State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China

^b Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing, China

^c School of Psychology, Beijing Normal University, Beijing, China

^d Department of Psychology and Social Behavior, University of California, Irvine, CA, United States

ARTICLE INFO

Article history:

Received 25 August 2014

Reviewed 2 October 2014

Revised 20 December 2014

Accepted 3 April 2015

Action editor Asaf Gilboa

Published online 23 April 2015

Keywords:

Subsequent memory

Spacing effect

Repetition effect

Storage strength

Retrieval strength

ABSTRACT

Although behavioral studies have consistently reported the spacing effect in learning, its cognitive and neural mechanisms are still not clearly elucidated. According to the storage/retrieval strength framework proposed by Bjork (1999; Bjork & Bjork, 1992), which was built on the study-phase retrieval hypothesis and the deficient processing hypothesis, the spacing effect is achieved by reducing memory retrieval strength during subsequent repetitions and thus enhancing storage strength. The present study tested this hypothesis with parallel fMRI and EEG. Participants were asked to study 180 words that were repeated three times, with half of the words having an inter-repetition-lag of 1–3 words (i.e., massed learning) and the other half an inter-repetition-lag of 25–35 words (i.e., spaced learning). An unexpected recognition test was administered 24 h after learning. Consistent with Bjork's hypothesis, the EEG data suggested that spaced learning was associated with weaker retrieval strength, as indicated by a reduced familiarity effect in frontal N400. Meanwhile, spaced learning effectively enhanced the encoding process and thus led to stronger storage strength, as indicated by greater fMRI responses during learning in brain regions whose activities were associated with subsequent memory. Interestingly, no direct association was found between repetition priming and episodic memory. These results furthered our understanding of the neural mechanisms underlying the spacing effect.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

First proposed by Ebbinghaus (1885/1964), the spacing effect is one of the most robust and consistent effects in learning and

memory (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Delaney, Verkoeijen, & Spiguel, 2010). It refers to the phenomenon that repeatedly studied items are better remembered when the repetitions are interleaved or distributed than when they are consecutive or massed (Greene, 1989;

* Corresponding author. State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, China.

E-mail address: guixue@gmail.com (G. Xue).

<http://dx.doi.org/10.1016/j.cortex.2015.04.002>

0010-9452/© 2015 Elsevier Ltd. All rights reserved.

Russo, Parkin, Taylor, & Wilks, 1998). Although the spacing effect has been observed across various learning materials (e.g., words, images, drawings, and motor skills) and tasks (e.g., recognition, free recall, and skill performance) (Bridger & Wilding, 2010; Cepeda et al., 2006; Goldstein, 1993; Marini, Marzi, & Viggiano, 2011; Otten, Sveen, & Quayle, 2007), its underlying cognitive and neural mechanisms are still under debate.

One of the most intuitive and widely accepted accounts for the spacing effect is the deficient processing hypothesis, which posits that compared to spaced presentation, massed presentation leads to less processing and thus worse memory. Deficient processing for massed learning can occur voluntarily when subjects pay less attention to the items they just studied, but it may also occur involuntarily. For example, a strong spacing effect was found even when subjects were required to pay full attention (i.e., to try not to reduce attention) to learning materials (D'Agostino & DeRemer, 1973; Elmes, Sanders, & Dovel, 1973; Hintzman, Summers, & Block, 1975). Similarly, when subjects are assumed not to voluntarily control their attention to what they are learning (e.g., incidental learning), researchers have observed the spacing effect (Challis, 1993; Reder, Park, & Kieffaber, 2009; Russo, Mammarella, & Avons, 2002). One mechanism that may lead to this involuntary effect is repetition priming/suppression. That is, the presentation of an item automatically primes and thus reduces the processing strength of its subsequent presentations. Compared with massed learning, spaced learning could reduce repetition priming and overcome deficient processing (Toppino, Fearnow-Kenney, Kiepert, & Teremula, 2009; Verkoeijen, 2005). The priming account is in line with evidence from many behavioral (Mammarella, Russo, & Avons, 2002; McKone, 1995; McKone & Dennis, 2000; Russo et al., 2002, 1998), functional MRI (Xue, Mei, et al., 2010; Xue et al., 2011) and EEG studies (Van Strien, Verkoeijen, Van der Meer, & Franken, 2007), which generally found greater behavioral repetition priming or neural repetition suppression under the massed condition than under the spaced condition.

Several methodological and theoretical issues deserve further examination. First, in several previous fMRI studies on the spacing effect, behavioral repetition priming was either not measured (Xue, Mei, et al., 2010; Xue et al., 2011) or measured between items (Wagner, Maril, & Schacter, 2000). Second, to increase the spacing effect, some studies used consecutive repetitions under the massed condition, which made the stimuli predictable and thus contaminated the encoding process (Xue, Mei, et al., 2010; Xue et al., 2011). Another study used an interval of one day for spaced learning (Wagner et al., 2000), which might have confounded the spacing effect with the effect of sleep-dependent memory consolidation. Third, when an intentional encoding task was used (Callan & Schweighofer, 2010), the results could be attributed to the effects of attention or voluntary verbal rehearsal. More importantly, the deficient processing hypothesis could not account for the inverted-U-shaped relationship between spacing and subsequent memory, that is, too long a repetition lag resulted in a lack of the spacing effect (Appleton-Knapp, Bjork, & Wickens, 2005).

The deficient processing hypothesis also suggests that episodic memory can be enhanced by reduced repetition

priming. Empirical support for this prediction has been mixed. Some studies found that items with greater behavioral or neural repetition priming were associated with worse episodic memory (Gagnepain, Lebreton, Desgranges, & Eustache, 2008; Wagner et al., 2000; Xue et al., 2011), but other studies found either no such relationship (Stark, Gordon, & Stark, 2008; Ward, Chun, & Kuhl, 2013; Xue et al., 2011) or even opposite relationship (Turk-Browne, Yi, & Chun, 2006).

In contrast to the deficient processing hypothesis, the storage and retrieval strength hypothesis of Bjork's new theory of disuse (NTD) (Bjork, 1999; Bjork & Bjork, 1992, 2006) does not have the above-mentioned limitations. According to Bjork, an item in memory can be characterized by two "strengths": Storage strength reflects how well an item is learned, and retrieval strength represents how accessible an item can be via recall or recognition. Whereas storage strength grows monotonically, retrieval strength decays as a function of time. Both storage and retrieval strengths can be increased through further learning and retrieval. Each time an item is studied again, it serves as a retrieval cue to reactivate the memory representation of the information stored during the first representation (Appleton-Knapp et al., 2005; Murray, 1983; Thios & Dagostino, 1976). The act of retrieval is itself a learning event in the sense that the retrieved information is more likely to be remembered than it would be without having been retrieved (Bjork, 1975, 1988). Importantly, increments in storage strength and retrieval strength are a decreasing function of the current retrieval strength because easily accessible items will not be deeply processed.

The NTD provides a good explanation not only for the spacing effect, but also for other effects related to repeated study that could not be explained by the deficient processing model. For the spacing effect, the NTD predicts that under the massed learning condition, items are readily accessible and retrieval strength is strong, but storage strength is increased to a lesser degree. Like the deficient processing hypothesis, the NTD suggests that spaced learning could overcome deficient processing as a result of repetition. It further specifies that the reduced processing strength is caused by the presence of momentary retrieval strength. Through the study-phase retrieval mechanisms, the NTD can explain an older finding that items that were not recognized at the second presentation were not recalled easily later on (Johnston & Uhl, 1976; Madigan, 1969). It can also explain the finding that too long a repetition lag resulted in a failure of retrieval, a lack of the spacing effect, and consequently an inverted-U-shaped relationship between spacing and subsequent memory (Appleton-Knapp et al., 2005). Finally, because there is a complex relationship between current retrieval/storage strength and further increase in retrieval/storage strength, and because repetition priming affects encoding strength, it would be difficult to establish a clear and direct relationship between repetition priming and subsequent memory.

Despite its strengths in explaining various behavioral results, the storage and retrieval strength model has not been examined with neuroimaging evidence. In particular, no neural evidence has been obtained to support the specific prediction that spacing is associated with a reduction in momentary retrieval strength, or an increase in encoding/storage strength. The present study tested these predictions

by using neural indices of encoding/storage and retrieval strength based on functional MRI and electrophysiological methods.

Encoding/storage strength can be indexed by activations during encoding in the task network, including the lateral prefrontal cortex, parietal lobule, the posterior material-specific processing areas, and the medial temporal lobe, because they have been found to be associated with successful subsequent memory (Kim, 2011; Liu, Dong, Chen, & Xue, 2014). For example, certain experimental conditions have led to both stronger activities in the above areas and better memory: deep encoding as compared to shallow encoding (Otten, Henson, & Rugg, 2001), focused attention as compared to divided attention (Uncapher & Rugg, 2005), spaced learning as compared to massed learning (Wagner et al., 2000; Xue, Mei, et al., 2010; Xue et al., 2011), and unfamiliar words as compared to familiar words (Chee, Westphal, Goh, Graham, & Song, 2003).

Retrieval strength can be indexed by ERP responses, particularly the frontal N400, which has been found to be sensitive to the familiarity effect during retrieval (Rugg & Curran, 2007; Rugg & Yonelinas, 2003). Specifically, the frontal N400 was more positive-going for recognized words than either new words or old words that were misjudged as new (Curran, 2000; Fernández et al., 1998; Padovani, Koenig, Brandeis, & Perrig, 2011; Rugg et al., 1998; Shimamura, 2011). This ERP component was also found to be more positive-going for massed repetitions than for spaced repetitions (Henson, Rylands, Ross, Vuilleumier, & Rugg, 2004; Kim, Kim, & Kwon, 2001; Van Strien et al., 2007). With the iEEG technique, Nahum et al. (2011) found that compared to spaced repetitions, massed repetitions induced a more positive ERP component between 250 and 400 msec located in the left MTL.

In this parallel fMRI and EEG study, participants were asked to study 180 words that were repeated three times. The inter-repetition-lag was 1–3 words for the massed learning condition, and 25–35 words for the spaced learning condition. An unexpected recognition test was administered 24 h after learning. Based on the above discussion about Bjork's storage/retrieval strength hypothesis and the likely neural indices of storage and retrieval strength, we proposed the following hypotheses. First, each repetition should increase momentary retrieval strength, but at the same time reduce processing strength and storage strength (Hypothesis 1, hereafter, H1). Second, given the inverted relationship between retrieval strength and further increase in storage and retrieval strength, we further predicted that the increase in retrieval strength and decrease in processing strength should be smaller from P2 to P3 than from P1 to P2 (H2). Third, compared to massed learning, spaced learning would reduce momentary retrieval strength, and at the same time would enhance encoding/storage strength and subsequent memory performance (H3). In addition to testing these three hypotheses from the NTD, we also examined the factors that might have affected repetition priming to understand the relationship between repetition priming and subsequent memory. Together, the current study provided an integral examination of the storage/retrieval strength hypothesis, and should advance our understanding of the neural mechanisms of the spacing effect.

2. Materials and methods

2.1. Participants

Twenty native Chinese volunteers (13 males, mean age = 19.6 years, ranging from 17 to 25 years) participated in the fMRI experiment and another 16 (7 males, mean age = 20.3 years, ranging from 19 to 22 years) participated in the ERP experiment. Data from three additional participants of the ERP experiment were excluded from further analysis because one subject had many (more than half) fast response trials (within 300 msec) and two other subjects had very few (fewer than 15) trials that were recognized with high confidence. All subjects were right-handed and had normal or corrected-to-normal vision. Informed consents were obtained before the experiment. This study was approved by the Institutional Review Board of the State Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University.

2.2. Stimuli

The same materials and design were used for both fMRI and EEG experiments. One hundred and eighty Chinese concrete nouns with medium to high word frequency were used as learning materials. Half of them describe living objects (e.g., elephant) and the other half nonliving objects (e.g., newspaper). The words used for spaced and massed learning were counterbalanced across subjects. To reduce the primacy and recency effects, 3 words were added at the beginning and the end of each run. These words were discarded in the following analysis. Another 180 nouns were used as foils in the recognition test. All the stimuli were presented in the same size (127*127 pixels) in white color on black background.

2.3. Procedure

A semantic judgment task was used during the encoding stage. Participants were asked to judge whether each word represented a living or nonliving object, by pressing a button with their left or right thumb (in the fMRI experiment) or index finger (in the ERP experiment). They were not told about, thus were not aware of, the subsequent memory test. Response keys were counterbalanced across participants. Each word was repeated three times, with half of the words under the massed learning condition with an inter-repetition-lag of 1–3 words, whereas the other half under the spaced learning condition with an inter-repetition-lag of 25–35 words (Fig. 1). For each trial, a stimulus was presented on the screen for 3 sec or until a response was made (in which case the word was replaced by a fixation), whichever came first. The next trial began after a jittered delay (0–5 sec). A fixation was shown on the screen between trials. Subjects finished three runs of the task, each lasting 12 min and 18 sec. The sequence was optimized for design efficiency for fMRI using an in-house Matlab script (Dale, 1999).

Twenty-four hours after the encoding task, participants were called back to the lab to complete an unexpected recognition task. The studied words and an equal number of foils were randomly mixed together. Participants were asked

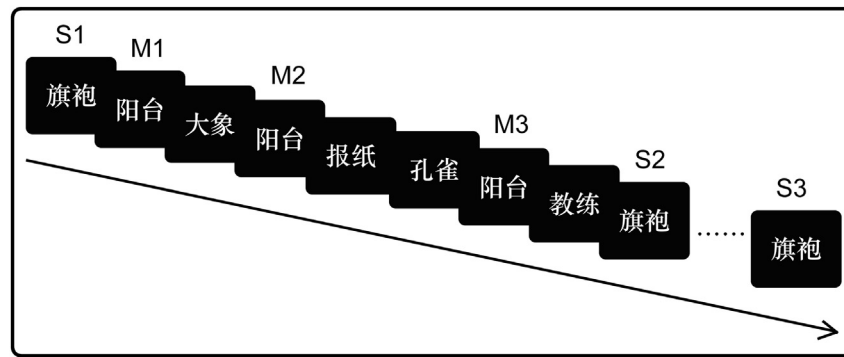


Fig. 1 – Experimental design. Subjects were asked to make living-nonliving judgment on each word that was repeated 3 times with an inter-repetition-lag of either 1–3 items (i.e., the massed condition) or 25–35 items (i.e., the spaced condition). In each trial, the word was presented on the screen for up to 3 sec, or until a response was made, whichever occurred first. Three words were added as fillers in the beginning and the end of each session to eliminate the primacy and recency effects. One day later, an unexpected recognition test was conducted to test memory performance. M: massed learning; S: spaced learning; the number followed them indicates the number of repetition. The English translations of Chinese words in the figures are, from left to right, cheongsam, balcony, elephant, balcony, newspaper, peacock, balcony, coach, cheongsam, cheongsam.

to decide whether the word presented on the screen was studied or not on a 6-point scale, with 1 = “Definitely new,” and 6 = “Definitely old.” There was no time pressure for the recognition task.

2.4. Behavioral data analysis

Accuracy and reaction times (RT) were calculated. For any trials whose RT exceeded two standard deviations, they and their repetitions were excluded from further analysis. In addition, trials containing EEG artifacts were also excluded (see the ERP data analysis section). As a result, 8.8% trials were excluded from the EEG analysis. No trials in the fMRI data needed to be excluded.

Preliminary analysis found no significant differences in accuracy or RT between the fMRI and the ERP experiments, so their behavioral data were analyzed together. We used RT differences between P1 and P2/P3 as the index of behavioral repetition priming using the following formula: $RP = P1 - (P2+P3)/2$. Memory performance was assessed with two indices. First, proportions of remembered and forgotten items were compared between the massed and the spaced conditions. To obtain comparable number of remembered and forgotten items for EEG and fMRI analysis, only the old words that were recognized as “Definitely old” (i.e., scored 6 on the 6-point scale) were defined as remembered items, and the old words scored 1–4 were defined as forgotten items. Items having a score of 5 were treated as a nuisance variable. Because the above index may be biased by participants’ response criteria, we further calculated the discriminability index (i.e., d') using the formula: $d' = Z_{(\text{hit rate})} - Z_{(\text{false alarm})}$ for each condition. To obtain the zROC curve, the standardized hit rates and false alarm rates were further fitted to a linear line with a least-squares model, separately for the spaced and massed conditions for each participant.

Correlational analysis (across items) was used to examine the relationships between RT at the first presentation

(hereafter P1 RT), repetition priming, spacing condition, and subsequent memory. The r values were obtained from each individual participant across trials, and then converted to Fisher’s z . T-tests were conducted to examine if they were significantly different from 0 at the group level. We also conducted a hierarchical logistic regression analysis to examine the independent contributions of P1 RT, repetition lag, repetition priming and semantic category to memory performance. A regression model was built for each participant. Group analysis was conducted on beta values using t-tests.

2.5. fMRI recording and data analysis

Functional MRI data were acquired on a 3.0-T Siemens MRI scanner in the MRI Center at Beijing Normal University. Participants lay supine on the scanner bed and viewed visual stimuli back-projected onto a screen through a mirror attached onto the head coil. Foam pads were used to minimize head motion. Single-shot T2*-weighted gradient-echo, EPI sequence was used for functional imaging acquisition with the following parameters: repetition time/echo time/ $\theta = 2000$ msec/30 msec/90°, field of view = 200 × 200 mm, matrix = 64 × 64, and slice thickness = 4 mm. Thirty contiguous axial slices parallel to AC–PC line were obtained to cover the whole cerebrum and partial cerebellum. Anatomical MRI was acquired using a T1-weighted, three-dimensional, gradient-echo pulse sequence. Parameters for this sequence were as follows: repetition time/echo time/ $\theta = 2530$ msec/3.39 msec/7°, field of view = 256 × 256 mm, matrix = 192 × 256, and slice thickness = 1.33 mm. One hundred twenty-eight sagittal slices were acquired to provide a high-resolution structural image of the whole brain.

Image preprocessing and statistical analysis were carried out using FEAT (fMRI Expert Analysis Tool) version 5.98, part of the FSL (FMRIB software library, version 4.1, www.fmrib.ox.ac.uk/fsl). The first three volumes before the task were

automatically discarded by the scanner to allow for T1 equilibrium. The remaining images were then realigned to correct for head movements (Jenkinson & Smith, 2001). Translational movement parameters never exceeded 1 voxel in any direction for any subject or session. Data were spatially smoothed using a 5-mm FWHM Gaussian kernel. The data were filtered in the temporal domain using a nonlinear high-pass filter with a 60-sec cutoff. A two-step registration procedure was used whereby EPI images were first registered to the MPRAGE structural image and then into the standard Montreal Neurological Institute (MNI) space using affine transformations (Jenkinson & Smith, 2001). Registration from MPRAGE structural image to the standard space was further refined using FNIRT nonlinear registration (Andersson, Jenkinson, & Smith, 2007; Andersson, Jenkinson, Smith, & Andersson, 2007). Statistical analyses were performed in the native image space, with the statistical maps normalized to the standard space prior to higher level analysis.

The general linear model within the FILM module of FSL was used to model the data, which examined the relations among spacing, repetition suppression, and subsequent memory. The words were separately modeled according to their memory performance (remembered versus forgotten), learning condition (massed versus spaced), and repetition (P1, P2 and P3). Only the old words that were recognized with high confidence (scored 6) were considered as remembered items (Otten et al., 2001). In contrast, the old words scored 1 to 4 were considered as forgotten items. Items scored 5 were treated as a nuisance variable. The first and the last three filler words were also treated as a separate nuisance variable. Null events were not explicitly modeled and therefore constituted an implicit baseline. For each subject, two contrast images were computed, including the main effects of subsequent memory (remembered versus forgotten) and spacing (spaced versus massed, using only P2 and P3 as there should be no differences in P1 response between the two conditions). A higher level analysis was conducted for cross-run contrasts for each subject using a fixed effect model. These were then entered into a random-effect model for group analysis, using FLAME (FMRIB's Local Analysis of Mixed Effects) Stage 1 only with automatic outlier detection (Beckmann, Jenkinson, & Smith, 2003; Woolrich, 2008; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004).

For this model, unless otherwise noted, group images were thresholded using cluster detection statistics, with a height threshold of $Z > 2.3$ and a cluster probability of $p < .05$, corrected for whole-brain multiple comparisons using Gaussian random field theory.

2.6. Conjunction analysis

To identify brain regions that showed both the subsequent memory effect (SME) and the spacing effect, a conjunction analysis was conducted using the procedure suggested by Nichols, Brett, Andersson, Wager, and Poline (2005). Group maps were thresholded at $Z = 2.3$ (cluster probability of $p < .05$) for the subsequent memory and spacing effects separately, and they were then binarized and multiplied.

2.7. ROI analysis

ROI analyses were conducted to examine the hypotheses regarding the encoding and storage strength. Clusters in the left fusiform cortex and left superior parietal lobule (SPL) that showed significant conjunctive effects were defined as ROIs. ROI analyses were performed by extracting parameter estimates (betas) of each event type from the fitted model and averaging them across all voxels in the cluster for each subject. Percent signal changes were calculated using the following formula: $[\text{contrast image}/(\text{mean of run})] \times \text{ppheight} \times 100\%$, where ppheight is the peak height of the hemodynamic response versus the baseline level of activity (Mumford, 2007). The difference between P1 and P2/P3 activity were used as the index of fMRI repetition suppression with the following formula: $RP = P1 - (P2 + P3)/2$.

2.8. EEG recording and data analysis

The ERP experiment was conducted in a soundproof, light adjustable room. Participants were seated comfortably about 100 cm away from the computer screen. EEG data were recorded from the scalp using a 64-channel NeuroScan system (Neurosoft, Inc., Sterling, VA). Ag – AgCl electrodes were mounted according to the international 10–20 system and all the electrodes were referenced to the left mastoid during on-line recording. The impedance of all electrodes was kept below 5 k Ω . The sampling rate was 1000 Hz.

EEGLAB (<http://www.sccn.ucsd.edu/eeglab/>) and in-house codes, running under the cross-platform MATLAB environment (the Mathworks, Inc. <http://www.mathworks.com/products/matlab>), were used to preprocess EEG data. Data were first downsampled to 500 Hz and filtered with band-pass of .1–40 Hz. The continuous EEG data from all channels were re-referenced to the global average reference (Lehmann & Skrandies, 1980). Eyeblink noise was identified and corrected using the independent component analysis algorithm in EEGLAB (Beres et al., 2002), and the components sensitive to eye blinks or movements were identified and removed manually. The continuous EEG data were segmented into epochs from 250 msec pre-stimulus until 1000 msec post-stimulus. The 250 msec pre-stimulus served as the baseline. Trials with artifacts exceeding $\pm 100 \mu\text{V}$ were excluded. Trials contaminated by residual eye movement, blinks, or amplifier saturation were rejected by visual inspection. The remaining trials were averaged, separately for each trial type and for each subject.

The trials were separated into 12 conditions according to their memory performance (remembered versus forgotten), learning condition (massed versus spaced), and repetition (P1, P2 and P3), which was the same as described above for the fMRI analysis. Previous studies suggest that frontal N400 component with a time window around 300–500 msec and maximal distribution in the frontal central part should be related to familiarity-based recognition, although the exact time window varied across studies (Van Strien, Hagenbeek, Stam, Rombouts, & Barkhof, 2005; Van Strien et al., 2007; Voss & Paller, 2009; Voss, Schendan, & Paller, 2010). In the present study, we defined the time window as 280–440 msec according to the visual inspection of the grand averaged EEG

waveforms, which showed a peak latency at 340 msec. The N400 amplitude was calculated by averaging the EEG response within this time window. Similar to the fMRI study, we first examined the SME and the spacing effect (using responses at P2 and P3). A conjunction analysis was then conducted to locate the electrodes showing both effects. Focusing on these electrodes and the N400 component, MANOVA was conducted to examine the specific hypotheses regarding retrieval strength. The N400 difference between P1 and P2/P3 were used as the index of ERP repetition priming with the following formula: $RP = P1 - (P2 + P3)/2$.

3. Results

3.1. Behavioral results

3.1.1. Spaced learning enhanced subsequent memory

Consistent with many previous observations, we found that spaced learning significantly increased the proportion of remembered items [52.83% in the massed condition versus 57.23% in the spaced condition, $t(35) = -4.25$, $p < .001$] and decreased the proportion of forgotten items [35.21% in the massed condition versus 31.29% in the spaced condition, $t(35) = 3.41$, $p = .0017$] (Fig. 2A). The d' under the spaced condition ($d' = 1.26$) was higher than that under the massed condition [$d' = 1.14$, $F(1, 35) = 20.58$, $p < .001$] (Fig. 2B). This result was further confirmed by the ROC analysis (Fig. 2C). Averaged across the participants, the massed and spaced conditions had similar slopes (.94 vs 1.00, $F < 1$), but different intercepts [1.18 vs 1.05, $F(1, 35) = 9.73$, $p = .0038$], suggesting

that spaced learning significantly improved the signal strength but not the standard deviation.

3.1.2. Spaced learning enhanced encoding strength by reducing repetition priming

Spacing (massed versus spaced) by repetition (P1 versus P2 versus P3) ANOVA on RT revealed a significant interaction [$F(2, 35) = 66.47$, $p < .0001$], suggesting that spaced learning reduced behavioral repetition priming (Fig. 2D). Further paired sample t -tests showed comparable P1 RT [712.6 msec vs 706.1 msec, $t(35) = 1.68$, $p = .102$], but much shorter RT for massed learning than for spaced learning at P2 [633.7 msec vs 673.6 msec, $t(35) = -11.04$, $p < .00001$], and P3 [607.3 msec vs 667.5 msec, $t(35) = -9.53$, $p < .0001$].

The accuracy was overall high (>95%) and further increased with repetitions [$F(2, 35) = 19.27$, $p < .0001$]. There was a marginal significant interaction between spacing and repetition [$F(2, 35) = 2.84$, $p = .065$] (Fig. 2E). Paired t -tests suggested the accuracy was comparable during P1 [96.82% vs 96.57%, $t(35) = .58$, $p = .56$], but lower for massed learning than for spaced learning at P2 [97.75% vs 98.49%, $t(35) = 3.15$, $p = .0034$], and P3 [97.44% vs 98.24%, $t(35) = -2.388$, $p = .0225$].

3.1.3. The relationship between P1 RT, repetition priming, spacing, and memory

Correlational analysis across items showed that there was a strong correlation between P1 RT and repetition priming [$r = .67$, $t(35) = 27.14$, $p < .0001$], suggesting that items with longer RT showed greater repetition priming. However, subsequent memory was not correlated with either P1 RT [$r = .02$, $t(35) = .96$, $p = .34$] or behavioral repetition priming [$r = .03$, $t(35) = 1.54$, $p = .13$].

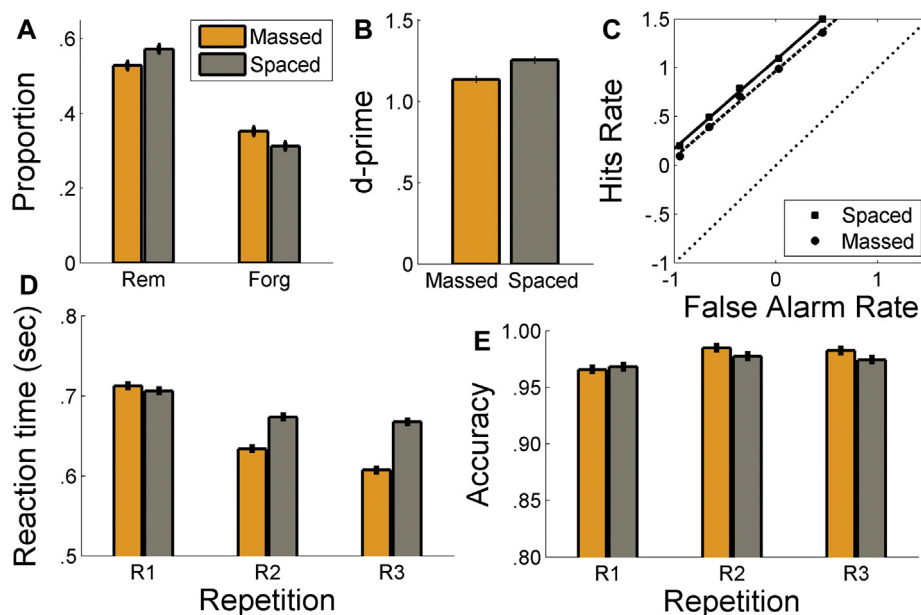


Fig. 2 – Behavioral results. Spaced learning enhanced memory performance as measured by both hit rates (A) and d' (B). Standardized hit rates and false alarm rates were further fitted to a linear line and the zROC curve (C) indicated that spacing increased the intercept but not the slope of the zROC curve. Reaction time (D) and accuracy (E) were plotted as a function of repetition and learning condition. Behavioral repetition priming was reduced under the spaced condition as compared to the massed condition. Error bars denote standard errors of the mean.

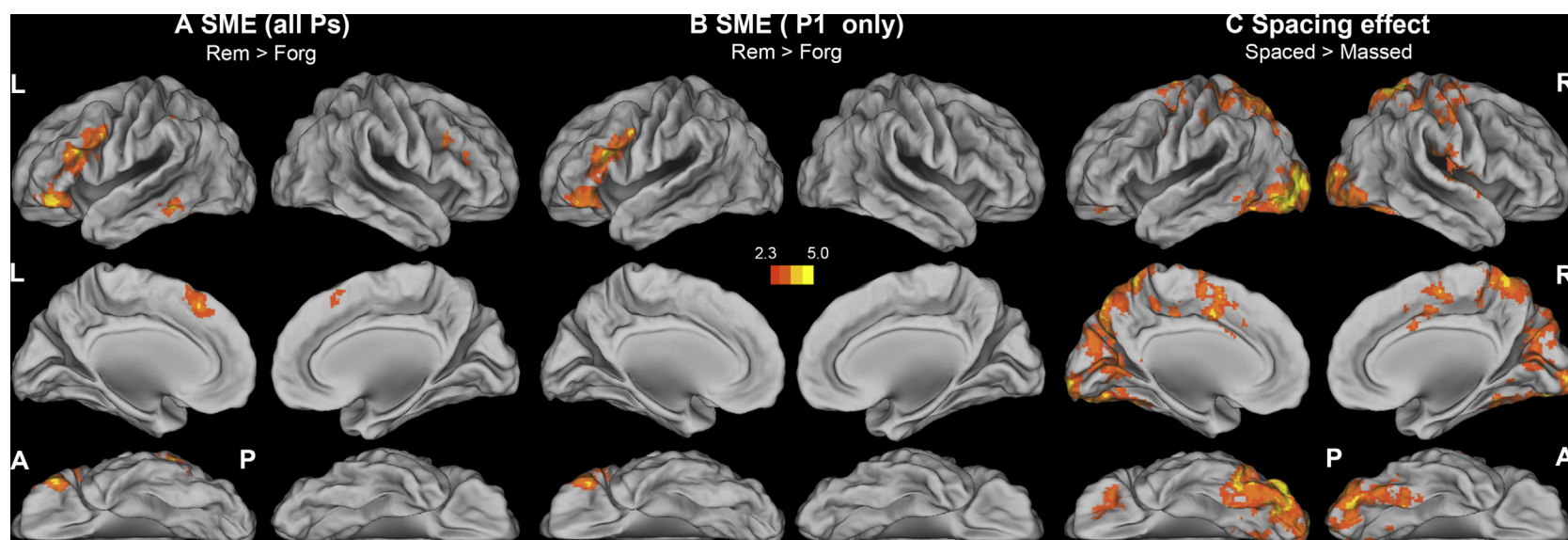


Fig. 3 – fMRI results. Brain regions showing the subsequent memory effect when using activity from all three repetitions (A), or when only using the P1 activity (B), and the spacing effect (C). All activations were thresholded at $z > 2.3$ (whole-brain corrected $p < .05$) and rendered onto a population-averaged surface atlas using multifiducial mapping (Van Essen, 2005). Rem: Remembered; Forg: Forgotten.

A logistic regression analysis was used to evaluate the unique contributions of P1 RT, repetition priming, and repetition lag to subsequent memory. Semantic category (living versus nonliving) was also included as a predictor because previous studies have shown an animacy effect in episodic memory (Bonin, Gelin, & Bugajska, 2013). The results showed that spaced learning [$\beta = -.25$, $t(35) = -4.43$, $p < .0001$] and living objects [$\beta = -.96$, $t(35) = -8.44$, $p < .0001$] were associated with better subsequent memory as compared with massed learning and nonliving objects, respectively. Notably, there was no effect of either behavioral repetition priming ($\beta = -.09$, $t < 1$) or P1 RT [$\beta = -.90$, $t(35) = -1.76$, $p = .09$].

Taken together, the behavioral analysis suggested that spaced learning enhanced subsequent memory. Repetition priming was affected by both P1 RT and spacing, but there was no strong relationship between reaction time and subsequent memory and no direct relationship between repetition priming and subsequent memory.

3.2. fMRI results

3.2.1. Brain regions associated with the spacing effect and SME

We first identified the brain regions showing a significant SME. A comparison of subsequently remembered items with forgotten items revealed significantly stronger activations for the former in the left superior parietal lobule (SPL, MNI: -34 , -60 , 48 , $Z = 3.69$), left fusiform cortex (MNI: -50 , -54 , -18 , $Z = 2.54$) that extended to the left inferior temporal gyrus, the left inferior frontal gyrus (IFG, MNI: -50 , 26 , 26 , $Z = 4.42$) that extended to the lateral orbitofrontal cortex (OFC, MNI: -46 , 42 , -8 , $Z = 4.24$), left superior frontal gyrus (SFG, MNI: -4 , 28 , 48 , $Z = 3.69$), and right middle frontal gyrus (MFG, MNI: 44 , 20 , 32 , $Z = 3.49$) (Fig. 3A, Table 1).

When only the P1 response was used to predict subsequent memory, we also found a significant SME in the left IFG (MNI: -48 , 24 , 22 , $Z = 4.64$) (Fig. 3B, Table 1). Apparently, this SME was weaker than when activities from all three repetitions were added together.

We then examined the brain regions that were modulated by spacing. Compared to massed learning, spaced learning increased brain activation in the left fusiform cortex (MNI: -46 , -56 , -20 , $Z = 4.67$) that extended to the left lateral occipital gyrus, left SPL (MNI: -48 , -28 , 48 , $Z = 3.7$), right frontal pole (MNI: 38 , 60 , 12 , $Z = 3.41$), and left OFC (MNI: -34 , 36 , -12 , $Z = 3.31$) (Fig. 3C, Table 1).

Conjunction analysis revealed that both the spacing effect and the SME were significant in three clusters: The left SPL (98 voxels, center of gravity in MNI: -28 , -66 , 36), left fusiform cortex (92 voxels, center of gravity in MNI: -50 , -56 , -22), and left OFC (30 voxels, center of gravity in MNI: -38 , 32 , -16) (Fig. 4A).

3.2.2. Examining the NTD hypotheses about encoding strength with fMRI data

Here we focused on the regions found in the conjunction analysis. Consistent with H1, three-way ANOVA revealed significant decreases in activity across repetitions in the left SPL [$F(2, 38) = 13.092$, $p < .0001$], left OFC [$F(2, 38) = 27.60$, $p < .0001$], and left fusiform [$F(2, 38) = 27.146$, $p < .0001$].

Table 1 – Brain regions showing P1 prediction of subsequent memory, the SME (remembered versus forgotten), and the spacing effect (spaced versus massed), as well as conjunction areas across the two effects.

Region	Volume (voxels)	Z	MNI coordinates		
			x	y	z
<i>P1 predicted memory</i>					
Left inferior frontal gyrus	2420	4.64	-48	24	22
<i>SME</i>					
Left inferior frontal gyrus	2717	4.42	-50	26	26
Left frontal orbital cortex		4.24	-46	42	-8
Left superior parietal lobule	522	3.69	-34	-60	48
Right middle frontal gyrus	450	3.49	44	20	32
Left superior frontal gyrus	366	3.69	-4	28	48
Left inferior temporal gyrus	342	4.15	-56	-44	-12
Left fusiform cortex		2.54	-50	-54	-18
<i>Spacing effect</i>					
Left lateral occipital cortex	30978	4.95	-36	-86	-10
Left fusiform cortex		4.67	-46	-56	-20
Right frontal pole	290	3.41	38	60	12
Left superior parietal lobule	283	3.7	-48	-28	48
Left frontal orbital cortex	273	3.31	-34	36	-12
<i>Conjunction areas</i>					
Left superior parietal lobule	98		-28	-66	36
Left fusiform cortex	92		-50	-56	-22
Left frontal orbital cortex	30		-38	32	-16

Notes: P1: first presentation; SME, subsequent memory effect.

Consistent with H3, there were significant interactions between spacing and repetition in these three regions [left SPL: $F(2,38) = 4.94$, $p = .012$]; the left OFC [$F(2, 38) = 3.812$, $p = .031$]; left fusiform: $F(2,38) = 4.90$, $p = .013$] (Fig. 4B). Further analysis found no significant difference between the spaced and massed conditions at P1 in the left SPL [$t(19) = 1.08$, $p = .295$], left OFC ($t < 1$), or left fusiform ($t < 1$). These results suggest that spacing reduced repetition suppression and thus enhanced the overall activity strength in regions that supported subsequent memory. Nevertheless, we did not find a significant interaction between the spacing effect and the SME in any of these three regions (left SPL & fusiform: $F_s < 1$; left OFC: $F(1, 19) = 2.625$, $p = .122$) (Fig. 4C), suggesting that the SME was comparable between the spaced and massed conditions.

To test H2, we compared the neural changes from P1 to P2 to those from P2 to P3. This analysis revealed significantly smaller changes from P2 to P3 than from P1 to P2 in the left SPL [$F(1,19) = 16.1$, $p = .0007$], left fusiform [$F(1,19) = 14.9$, $p = .0011$], and left OFC [$F(1,19) = 14.6$, $p = .0011$], for both spaced and massed conditions.

3.2.3. No interactions between repetition suppression and subsequent memory

The whole brain analysis revealed no regions showing significant interactions between repetition suppression and subsequent memory. When focusing on the three regions showing a significant conjunctive effect, three-way ANOVA still revealed no interaction between subsequent memory and neural repetition suppression [$F(2,38) = 1.35$, $p = .27$ in the left SPL; $F(2,38) = 1.90$, $p = .163$ in the left OFC; $F(2,38) = 1.94$,

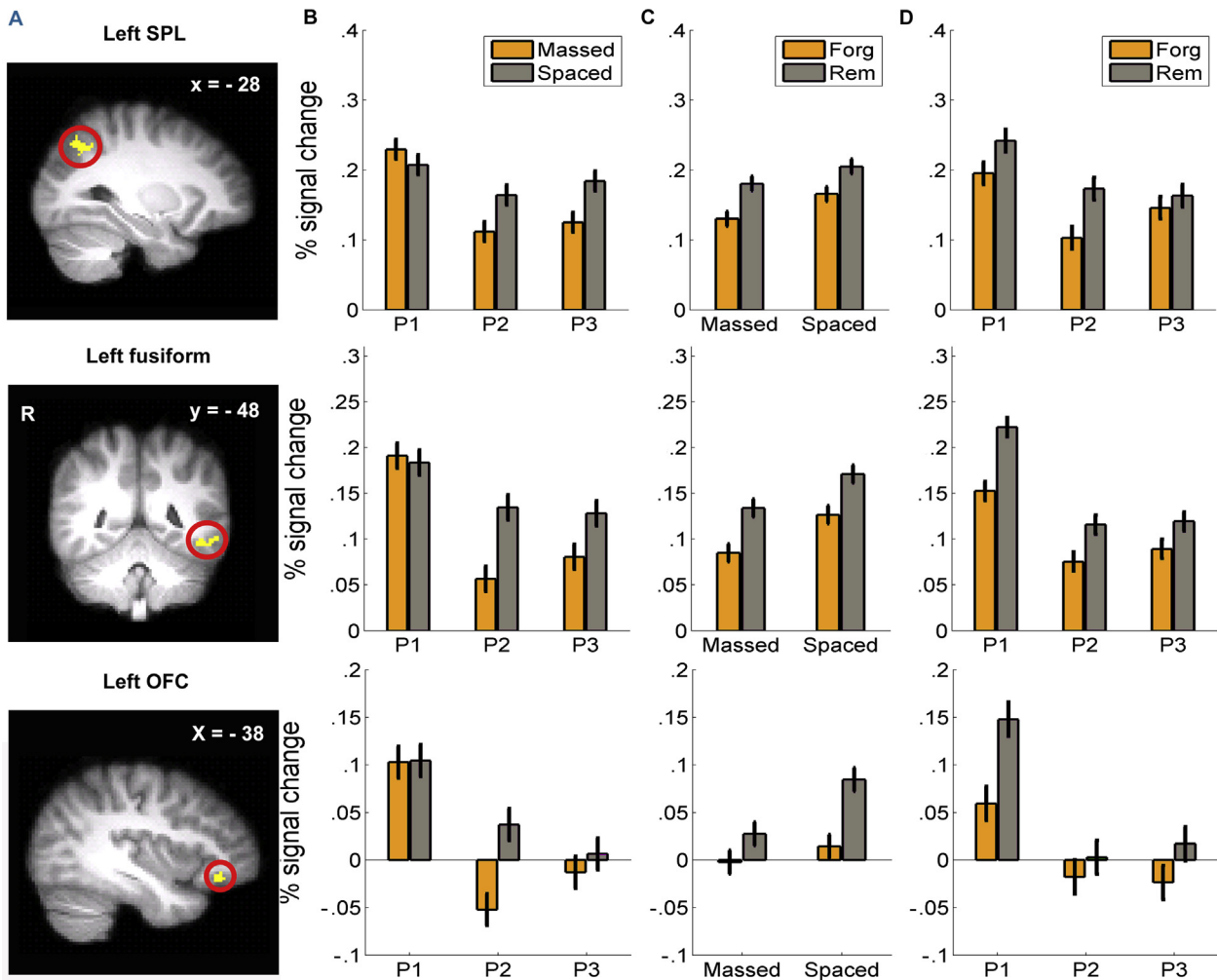


Fig. 4 – Conjunctural results. A. Regions showing both subsequent memory and spacing effects in the left SPL, left fusiform, and left OFC. Results were overlaid on the sagittal (upper and lower panels) and coronal (middle panel) slices of the group mean structural image. B, C and D show the plots of % signal change in these regions, as a function of repetition, learning condition, and subsequent memory performance. Error bars denote the within-subject standard errors.

$p = .16$ in the left fusiform] (Fig. 4D). Finally, we examined the interactions separately for the spaced and massed conditions to see if there was any evidence of such interactions. This analysis revealed no significant effects (all $ps > .183$). In sum, there was no evidence of direct interactions between neural repetition suppression and subsequent memory.

Consistent with the behavioral data on repetition priming, we found that repetition suppression was affected by both P1 response and spacing (Fig. 5 A–C). In particular, brain activation for P1 was significantly and positively correlated with neural repetition suppression under both massed ($r = .73$, $p < .001$ in the left SPL; $r = .81$, $p < .0001$ in the OFC; $r = .81$, $p < .0001$ in the left fusiform) and spaced conditions ($r = .52$, $p = .019$ in the left SPL; $r = .78$, $p < .0001$ in the OFC; $r = .53$, $p = .017$ in left fusiform).

Finally, we found that the relationship between behavioral repetition priming and neural repetition suppression was not robust or consistent (Fig. 5 D–F). Under the spaced condition, the correlation was not significant in either the left OFC ($r = -.06$, $p = .808$) or the left fusiform ($r = .29$,

$p = .211$), and only marginally significant in the left SPL ($r = .42$, $p = .068$). Under the massed condition, the correlation was significant in the left OFC ($r = .50$, $p = .026$), but not significant in the left fusiform ($r = .40$, $p = .083$) or SPL ($r = -.006$, $p = .98$).

3.2.4. The effect of processing time on the BOLD response

In the present study, we found that there was a significant difference in the RT for different conditions. Because word presentation ended as soon as the participant responded, the duration of stimulus presentation differed by condition, which would lead to possible differences in perceptual processes and BOLD response between conditions. Although we think the RT difference could be an inherent contributor to the spacing effects, it would be informative to examine the RT effect on brain responses. To this end, we tested a new model that included RT as a covariate. This analysis revealed that the major results remained unchanged although there were some minor changes (see [Supplementary Materials for details](#)).

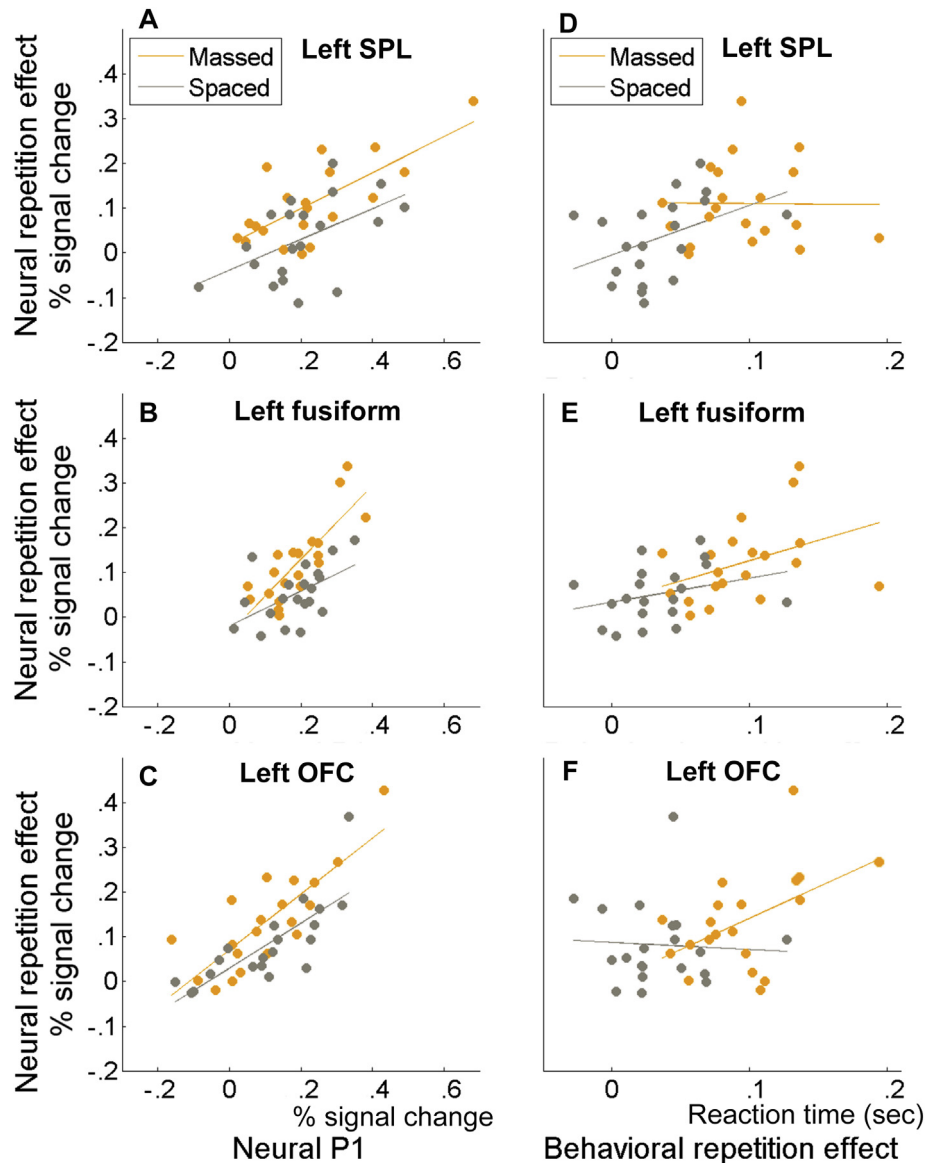


Fig. 5 – Correlation analysis. Correlations between neural repetition suppression and brain activity on P1 were significant under both massed and spaced conditions in the left SPL (A), left fusiform (B), and left IFG (C). The x-axis represents the fMRI activation in P1 and y-axis represents repetition suppression calculating fMRI activation difference $RP = P1 - (P2 + P3)/2$. However, the correlation between neural repetition suppression and behavioral repetition priming was significant only in the IFG (F), but not significant or only marginally significant in the SPL (D) and fusiform (E). The x-axis represents the behavioral repetition priming calculated as RT difference $RP = P1 - (P2 + P3)/2$ and y-axis represents repetition suppression calculating fMRI activation difference $RP = P1 - (P2 + P3)/2$.

3.3. ERP results

3.3.1. The subsequent memory and spacing effects on N400
 Consistent with previous studies (Marini et al., 2011; Van Strien et al., 2007; Voss & Paller, 2009), results revealed a more positive going frontal N400 for subsequently remembered than forgotten items in the frontocentral electrodes (Fig. 6A). Meanwhile, compared with spaced learning, massed learning was associated with a more positive going frontal N400 in similar locations. Conjunction analysis revealed that 11 electrodes (F1, FZ, F2, F4, FC1, FCz, FC2, FC4,

C1, Cz, C2) showed common effects of spacing and subsequent memory.

3.3.2. Examining the NTD hypotheses regarding retrieval strength with N400

Focusing on the frontocentral electrodes and the N400 response, three-way MANOVA revealed that consistent with H1, there was a significant repetition priming effect [$F(2, 30) = 3.23, p = .05$], suggesting that repetition increased momentary retrieval strength as indicated by the more positive going N400.

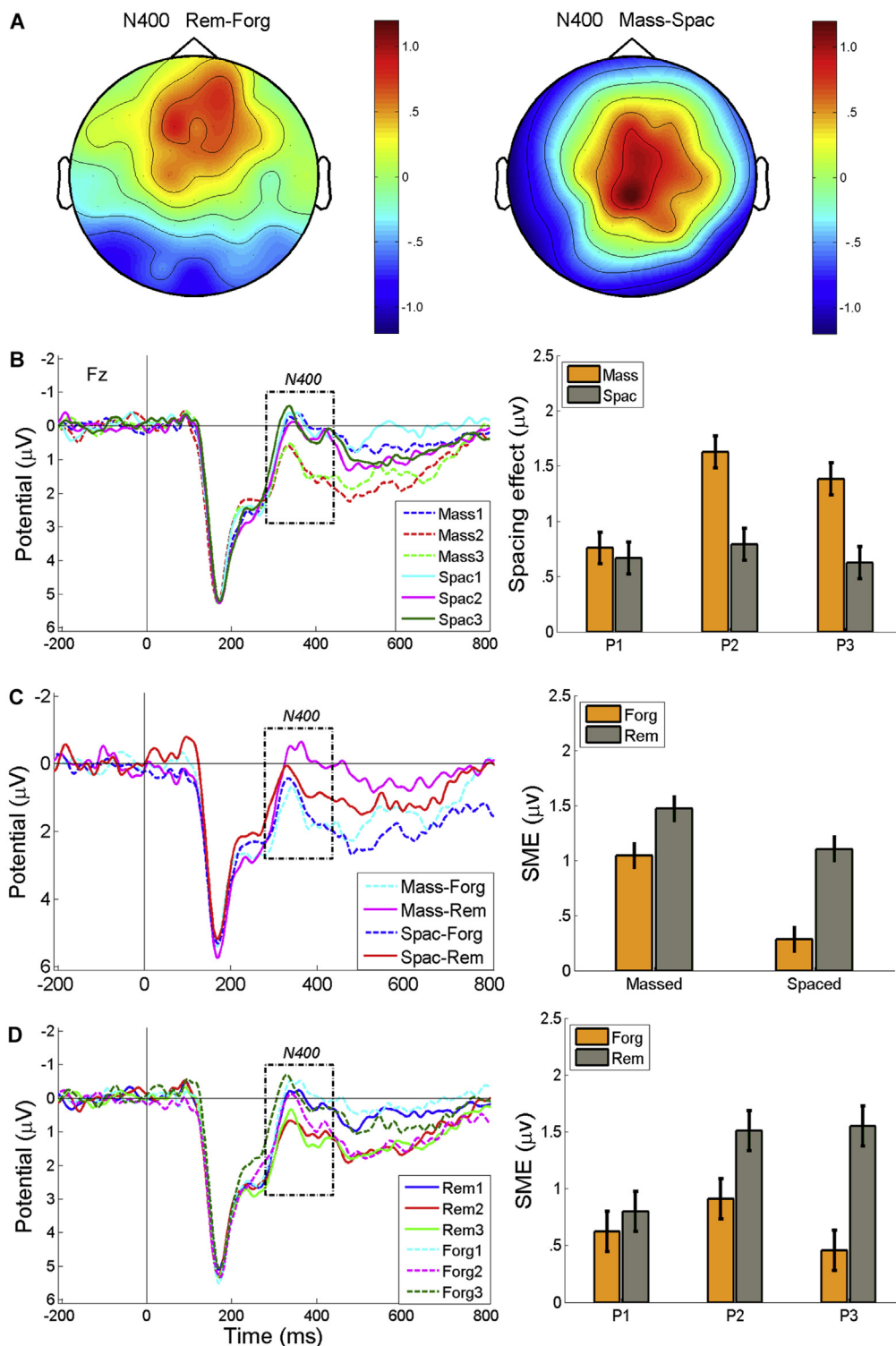


Fig. 6 – EEG Results. The subsequent memory effect (A, left panel) and the spacing effect (A, right panel) shown on the topographic maps. B, C, and D show the representative EEG waveforms in the Fz electrode (left panel), and the mean frontal N400 amplitudes (right panel), as a function of repetition, learning condition, and subsequent memory performance of spacing, repetition, and/or subsequent memory. Mass = Massed; Spac = Spaced.

Also consistent with H3, we found a significant spacing-by-repetition interaction [$F(2, 15) = 3.97, p = .030$] (Fig. 6B). Post hoc paired sample *t*-tests showed that the amplitudes of the frontal N400 for the two conditions were comparable at P1 (.76 vs. .66, $t < 1$), but more positive for the massed condition than for the spaced condition at P3 [1.38 vs. .62, $t(15) = 3.44, p = .0036$]. Again, we did not find a significant interaction between spacing and subsequent memory [$F(1, 15) = 3.32, p = .088$] (Fig. 6C).

To test H2, we compared the neural changes from P1 to P2 to those from P2 to P3. This analysis revealed marginally smaller changes in the N400 component [$F(1,15) = 4.25, p = .057$] from P2 to P3 than from P1 to P2, for both spaced and massed conditions.

3.3.3. The interaction between repetition priming and subsequent memory

Unlike the fMRI results, we found that the interaction between repetition and subsequent memory was significant [$F(2, 30) = 3.39, p = .047$] (Fig. 6D). Post hoc paired sample *t*-tests showed that the frontal N400 amplitude was comparable at P1 [.80 vs. .62, $t(1,15) = .88, p = .39$], but was more positive for subsequently remembered items than forgotten items at P3 [1.55 vs. .46, $t(1,15) = 6.30, p < .0001$].

Correlational analysis showed that P1 response (N400) was not correlated with the neural repetition effect either under the massed ($r = -.008, p = .98$) or the spaced condition ($r = .34, p = .20$). Meanwhile, the correlation between behavioral repetition priming and the neural (N400) repetition effect was not significant under either the massed ($r = .16, p = .56$) or the spaced condition ($r = .07, p = .80$).

4. Discussion

In this study, we investigated the behavioral and neural mechanisms of the spacing effect in episodic memory. Consistent with the storage/retrieval strength hypothesis of the NTD (Bjork, 1999; Bjork & Bjork, 1992, 2006), which posits that high retrieval strength impairs the enhancement of long-term storage strength, we found that repeated studies were associated with increased retrieval strength, which impaired further enhancement of storage and retrieval strength. Spaced learning, on the other hand, could reduce momentary retrieval strength during subsequent repetitions, thus enhanced storage and retrieval strength and memory. In addition, our data help us to understand the complex relationship between repetition priming/suppression and subsequent memory, and thus shed light on the interaction between implicit and explicit memory.

The primary prediction of the NTD is that repeated studies increase retrieval strength, but impair further enhancement of storage and retrieval strength. In other words, increments in storage strength and retrieval strength are a decreasing function of the current retrieval strength. Our data are quite consistent with this prediction. We found a more positive going N400 with each repetition, accompanied by decreased reaction time and reduced BOLD responses in the brain regions that support memory encoding. Critically, we found a big increase in retrieval strength (as reflected by N400) from P1

to P2, but no further increase from P2 to P3. Correspondingly, we observed a big decrease in encoding strength from P1 to P2, but no further decrease from P2 to P3.

Similar to the deficient processing hypothesis, the NTD predicts that spaced learning would be associated with stronger encoding and better memory. This has been supported by many behavioral and functional imaging studies (Henson et al., 2004; Mei et al., 2010; Stark et al., 2008; Toppino et al., 2009; Van Strien et al., 2007; Verkoeijen, 2005; Wagner et al., 2000; Xue, Mei, et al., 2010; Xue et al., 2011). Consistent with these observations, we found that better memory was achieved by enhancing the activities in regions responsible for subsequent memory. Furthermore, we found that these enhanced responses were achieved by reducing repetition priming/suppression. It should be noted that although the BOLD responses in the current studies were considered as an index of processing strength that helps to further increase storage and retrieval strength, we did not suggest they were involved in memory encoding *per se*. Indeed, these activations were partly caused by the demand for retrieval, such as semantic retrieval and study-phase retrieval during subsequent repetitions. According to Bjork's storage/retrieval strength framework in NTD, the retrieval process is critical for memory encoding, reflecting a tight interaction between encoding and retrieval.

The current study had three major improvements over previous studies that showed similar results. First, both behavioral repetition priming and neural repetition suppression were directly measured across repetitions, which allowed for a more accurate estimation of the repetition effect. Second, the three repetitions under the massed condition were not consecutive, which not only provided a stronger test of the spacing effect, but also eliminated the contamination of predictability on the encoding process (Xue, Mei, et al., 2010; Xue et al., 2011). Third, we used an incidental encoding task (semantic judgment) rather than an intentional encoding task, so our results could not be attributed to the effects of voluntary verbal rehearsal (Callan & Schweighofer, 2010). Therefore, our behavioral and neural data together provide strong evidence that spacing enhances memory by reducing repetition priming/suppression.

Critically, with the identical design, our EEG results complemented the fMRI results and provided further support to the additional hypothesis proposed by the NTD: Spacing enhances memory by impairing momentary retrieval strength. We found that, consistent with previous studies, subsequently remembered items showed a greater frontal N400 familiarity effect (Rugg & Curran, 2007; Rugg et al., 1998; Rugg & Yonelinas, 2003). We also found that repetitions were associated with a more positive going frontal N400, which was attenuated by spacing (Henson et al., 2004; Kim et al., 2001; Nahum et al., 2011). Meanwhile, a recent study of the spacing effect on recognition (Ferrari, Bradley, Codispoti, Karlsson, & Lang, 2013) found that during the final test, the parietal N400 was greatest for distributed words, followed by massed words, single-presentation words, and new words, suggesting that practice enhanced the final storage strength and that spaced repetitions were more effective than massed repetitions. Taken together, these studies suggest that spacing can reduce momentary retrieval strength during encoding but

enhance storage strength, providing support to the storage/retrieval strength hypothesis.

By using a continuous recognition task for encoding, a previous EEG study also found that repeated words were associated with a stronger fronto-parietal N400, and that the increase in N400 was bigger for massed repetitions than for spaced repetitions (Van Strien et al., 2007). Van Strien et al. interpreted the more positive going N400 as reflecting weaker semantic processing, in support of the deficient processing hypothesis. However, they did not examine the SME. Our finding that subsequently remembered items showing a more positive N400 did not seem to fit the semantic encoding account, but instead favored the familiarity/retrieval strength account.

The above evidence suggests that the frontal BOLD response and the EEG N400 might reflect different aspects of memory processing, with the BOLD responses reflecting more of processing strength and the EEG more of retrieving strength. Consistently, previous studies suggested that the BOLD signals (Otten et al., 2001), but not the N400 (Rugg et al., 1998; Rugg & Yonelinas, 2003) was affected by the depth of processing. Meanwhile, the frontal N400 was more positive going when retrieval familiar words or massed words (Rugg & Curran, 2007; Rugg & Yonelinas, 2003), but the BOLD activation was stronger for unfamiliar words or spaced words which need more effort of encoding or processing (Chee et al., 2003; Wagner et al., 2000; Xue, Mei, et al., 2010; Xue et al., 2011). Nevertheless, this interpretation could be over simplified, and future studies should examine the underlying cognitive processes reflected by fMRI BOLD and EEG responses.

Our results also shed light on the relationship among spacing, repetition priming/suppression, and episodic memory. Although episodic memory and repetition priming have been traditionally considered as two distinct memory systems and supported by non-overlapping neural systems (Levy, Stark, & Squire, 2004; Nissen, Knopman, & Schacter, 1987; Reder et al., 2009; Schott et al., 2006; Squire, 2004; Ward et al., 2013), cumulative research has shown that these two types of memories involve at least partially overlapping mechanisms. For example, attention (Turk-Browne et al., 2006; Yi & Chun, 2005), depth of encoding (Otten et al., 2001), spacing (Wagner et al., 2000; Xue et al., 2011), and other manipulations can simultaneously affect priming and episodic memory (Paller, Voss, & Boehm, 2007; Richardson-Klavehn & Bjork, 1988; Yonelinas, 2002). Consistently, computational models suggest that the apparent dissociation between priming and episodic memory can be accounted for by common encoding processes but differential retrieval processes (Berry, Shanks, & Henson, 2008).

In a series of studies, Paller and colleagues (Paller et al., 2007) have suggested that the frontal N400, traditionally associated with familiarity, might also reflect conceptual priming. In their studies, conceptual priming was manipulated by including conceptually primed versus unprimed faces (Voss & Paller, 2006) or high- versus low-meaning “squiggle” stimuli (Voss & Paller, 2007). They consistently found a strong frontal N400 effect for repetition priming. In the present study, repetition priming/suppression was manipulated by spacing, and our results showed that the frontal N400 was modulated by both semantic priming and spacing, consistent with its role

in semantic priming. Furthermore, we found that the frontal N400 was predictive of subsequent memory, consistent with its role in familiarity. Given that the frontal N400 can index both repetition priming and subsequent memory, and spacing can affect both, our results suggest the N400 might reflect some common processes that support both repetition priming and episodic memory.

The storage/retrieval strength hypothesis of the NTD further suggests that the observed repetition priming/suppression could be attributed to momentary retrieval strength. Our results provide two pieces of evidence for this conjecture. First, we found that repetition priming was affected by initial encoding strength, as reflected in the positive correlations between P1 RT and behavioral repetition priming and between P1 response and fMRI repetition suppression. Second, behavioral, fMRI and EEG evidence together suggested that spacing reduced momentary retrieval strength and repetition priming and suppression. Together, the existing evidence suggests that repetition priming and episodic memory are more connected than previously thought.

Because of the complex relationship between the current retrieval strength and further increase of retrieval and storage strength, the NTD would not predict a clear association between repetition priming and subsequent memory. This is indeed the case. Adding to the rather mixed empirical observations (see Introduction), the current study found little evidence for the interaction between repetition priming and subsequent memory. Our results seemed to provide some clues to the elusive link between the repetition effect and the SME. First, reaction time might not be a pure indicator of cognitive and neural processes underlying memory encoding. For example, although brain activity is consistently predictive of subsequent memory, reaction time is not, as shown by the current and previous studies (Liu et al., 2014). Meanwhile, behavioral repetition priming could reflect multiple components of learning, such as stimulus, decision, and response (Dobbins, Schnyer, Verfaellie, & Schacter, 2004; Horner & Henson, 2008; Race, Shanker, & Wagner, 2009), and its correlation with neural repetition priming depends critically on the nature of these processes (Horner & Henson, 2008; Race et al., 2009). Consistently, the current study found that behavioral repetition priming was correlated with activity in the IFG but not in the SPL or fusiform. As behavioral measures could not help to dissociate these various components in order to examine their respective contributions to memory encoding, it is difficult to establish quantitative relationship between repetition priming and subsequent memory.

Second, both behavioral repetition priming and neural repetition suppression are affected by multiple factors such as P1 RT/neural activity and spacing. As pointed out by Xue et al. (2011), P1 response, spacing, and repetition priming have contradictive effects on storage strength and subsequent memory. Although the NTD provides qualitative hypotheses regarding how multiple learning events are integrated to enhance storage and retrieval strength, a quantitative model is still missing. The present study and several previous studies suggest that the summed activities across repetitions are the strongest predictor of the SME (compared to P1 activity and repetition suppression). These summed activities might only represent approximations rather than quantitative measures

of the overall processing strength and storage strength. For example, our previous study showed that the same brain activity level was associated with remembering under the massed condition but with forgetting under the spaced condition, suggesting at least some brain activities under the spaced condition did not convert to storage strength (Xue et al., 2011).

How to resolve this issue remains a great challenge for researchers in this field. Yet it has implications for a more fundamental question regarding how multiple presentations are summed to form better/stronger memories (Xue, Dong, et al., 2010). One potentially useful approach is to control the variability of material to match the P1 response, thus all differences in processing strength could be captured by repetition suppression. For example, when the P1 response for subsequently remembered and forgotten items was matched, we found significant or marginally significant interactions between repetition suppression and subsequent memory in a previous fMRI study (Xue et al., 2011) and in the current EEG experiment.

Unlike the continuous recognition task, subjects in the current study were not required to explicitly retrieve previous studies upon subsequent repetition. Did study-phase retrieval still occur? Behavioral and computational studies have consistently suggested that during repeated studies, study-phase retrieval can benefit memory encoding (Appleton-Knapp et al., 2005; Thios & Dagostino, 1976). Imaging studies suggest that study-phase retrieval is accompanied by the reactivation of early neural activation pattern (Kuhl, Shah, DuBrow, & Wagner, 2010; Lu, Wang, Chen, & Xue, 2015; Xue et al., 2013). This item-specific activation reinstatement precedes memory (Polyn, Natu, Cohen, & Norman, 2005) and is associated with performance in free recall (Gelbard-Sagiv, Mukamel, Harel, Malach, & Fried, 2008) and cued-retrieval (Kuhl, Rissman, Chun, & Wagner, 2011). In addition, reinstatement is not specific to an individual item but also its temporal context (Manning, Polyn, Baltuch, Litt, & Kahana, 2011; Miller et al., 2013).

Several important questions remain to be answered. First, although we used both EEG and fMRI techniques to study the relationship between repetition suppression and subsequent memory, we did not record both signals simultaneously. As suggested by the current study and several previous studies, the repetition effect observed in behavior, ERP, and fMRI data may reflect at least partially non-overlapping cognitive processes (Henson et al., 2004; Raaijmakers, 2003). Simultaneous EEG and fMRI would be useful to link familiarity with processing strength, which may provide a stronger test of the retrieval/storage strength hypothesis. In particular, the NTD specifically predicts that there is an interaction between items' current retrieval/storage strength and further gains of both storage and retrieval strength (Bjork, 1999; Bjork, 1992, 2006). Simultaneous EEG and fMRI recording would help to test this hypothesis. Second, although retrieval has been proposed as a memory modifier and its effect has been supported by many behavioral studies, its exact neural mechanisms remain to be examined to deepen our insights into the spacing effect. In particular, future studies should examine how different types of retrieval (study-phase retrieval versus continuous recognition) and different outcomes of

retrieval (familiarity versus recollection) affect the further increase of retrieval and storage strength. Finally, as mentioned earlier, future studies need to develop a quantitative model to describe the relationship between brain activities and storage and retrieval strength, and how multiple learning events are integrated to form long-lasting memory. Such a model would significantly advance our understanding of how repeated practices enhance memory, a well-known phenomenon but whose mechanisms are still poorly understood.

Acknowledgments

This work was partly sponsored by the National Natural Science Foundation of China (31130025), The 973 Program (2014CB846102), the 111 Project (B07008), and the NSFC Project (31221003).

Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.cortex.2015.04.002>.

REFERENCES

- Andersson, J. L., Jenkinson, M., & Smith, S. (2007a). *Non-linear registration. aka Spatial normalisation* FMRIB technical report TR07JA2. FMRIB Analysis Group of the University of Oxford.
- Andersson, J. L., Jenkinson, M., Smith, S., & Andersson, J. (2007b). *Non-linear optimisation*. FMRIB technical report TR07JA1. Oxford (UK): FMRIB Centre.
- Appleton-Knapp, S. L., Bjork, R. A., & Wickens, T. D. (2005). Examining the spacing effect in advertising: encoding variability, retrieval processes, and their interaction. *Journal of Consumer Research*, 32(2), 266–276.
- Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in FMRI. *NeuroImage*, 20(2), 1052–1063.
- Beres, S. B., Sylva, G. L., Barbian, K. D., Lei, B., Hoff, J. S., Mammarella, N. D., et al. (2002). Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proceedings of the National Academy of Sciences*, 99(15), 10078–10083.
- Berry, C. J., Shanks, D. R., & Henson, R. N. (2008). A unitary signal-detection model of implicit and explicit memory. *Trends in Cognitive Sciences*, 12(10), 367–373.
- Bjork, R. (1975). Retrieval as a memory modifier. In M. M. Grunberg, P. E. Morris, & R. N. Sykes (Eds.), *Information processing and cognition: The Loyola Symposium* (pp. 396–401). London: Wiley.
- Bjork, R. (1988). Retrieval practice and the maintenance of knowledge. In R. L. Solso (Ed.), *Practical aspects of memory II* (pp. 123–144). Hillsdale, NJ: Erlbaum.
- Bjork, R. A. (1999). Assessing our own competence: heuristics and illusions. In D. Gopher, & A. Koriati (Eds.), *Attention and performance* (pp. 435–459). Cambridge, MA, US: The MIT Press.
- Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning processes to cognitive processes:*

- Essays in honor of William K. Estes (Vol. 2, pp. 35–67). Hillsdale, NJ: Erlbaum.
- Bjork, R. A., & Bjork, E. L. (2006). Optimizing treatment and instruction: implications of a new theory of disuse. *Memory and society: Psychological perspectives*, 116–140.
- Bonin, P., Gelin, M., & Bugaiska, A. (2013). Animates are better remembered than inanimates: further evidence from word and picture stimuli. *Memory & Cognition*, 1–13.
- Bridger, E. K., & Wilding, E. L. (2010). Requirements at retrieval modulate subsequent memory effects: an event-related potential study. *Cognitive Neuroscience*, 1(4), 254–260.
- Callan, D. E., & Schweighofer, N. (2010). Neural correlates of the spacing effect in explicit verbal semantic encoding support the deficient-processing theory. *Human Brain Mapping*, 31(4), 645–659.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354.
- Challis, B. H. (1993). Spacing effects on cued-memory tests depend on level of processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(2), 389.
- Chee, M. W., Westphal, C., Goh, J., Graham, S., & Song, A. W. (2003). Word frequency and subsequent memory effects studied using event-related fMRI. *NeuroImage*, 20(2), 1042–1051.
- Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, 28(6), 923–938.
- D'Agostino, P. R., & DeRemer, P. (1973). Repetition effects as a function of rehearsal and encoding variability. *Journal of Verbal Learning and Verbal Behavior*, 12(1), 108–113.
- Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Hum Brain Mapp*, 8(2–3), 109–114.
- Delaney, P. F., Verkoeijen, P., & Spigel, A. (2010). Spacing and testing effects: a deeply critical, lengthy, and at times discursive review of the literature. In B. H. Ross (Ed.), *Psychology of learning and motivation: Advances in research and theory* (Vol. 53, pp. 63–147). San Diego: Elsevier Academic Press Inc.
- Dobbins, I. G., Schnyer, D. M., Verfaellie, M., & Schacter, D. L. (2004). Cortical activity reductions during repetition priming can result from rapid response learning. *Nature*, 428(6980), 316–319.
- Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). Leipzig: Von Duckner and Humblot (Reprinted from: Dover Publications).
- Elmes, D. G., Sanders, L. W., & Dovel, J. C. (1973). Isolation of massed and distributed-practice items. *Memory & Cognition*, 1(1), 77–79.
- Fernández, G., Weyerts, H., Schrader-Bölsche, M., Tendolkar, I., Smid, H. G., Tempelmann, C., et al. (1998). Successful verbal encoding into episodic memory engages the posterior hippocampus: a parametrically analyzed functional magnetic resonance imaging study. *The Journal of Neuroscience*, 18(5), 1841–1847.
- Ferrari, V., Bradley, M. M., Codispoti, M., Karlsson, M., & Lang, P. J. (2013). Repetition and brain potentials when recognizing natural scenes: task and emotion differences. *Social Cognitive and Affective Neuroscience*, 8(8), 847–854.
- Gagnepain, P., Lebreton, K., Desgranges, B., & Eustache, F. (2008). Perceptual priming enhances the creation of new episodic memories. *Consciousness and Cognition*, 17(1), 276–287.
- Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., & Fried, I. (2008). Internally generated reactivation of single neurons in human hippocampus during free recall. *Science*, 322(5898), 96–101.
- Goldstein, I. L. (1993). *Training in organizations: Needs assessment, development, and evaluation*. Thomson Brooks/Cole Publishing Co.
- Greene, R. L. (1989). Spacing effects in memory: evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(3), 371.
- Henson, R., Rylands, A., Ross, E., Vuilleumier, P., & Rugg, M. (2004). The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *NeuroImage*, 21(4), 1674–1689.
- Hintzman, D. L., Summers, J. J., & Block, R. A. (1975). Spacing judgments as an index of study-phase retrieval. *Journal of Experimental Psychology: Human Learning and Memory*, 1(1), 31.
- Horner, A. J., & Henson, R. N. (2008). Priming, response learning and repetition suppression. *Neuropsychologia*, 46(7), 1979–1991.
- Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical image analysis*, 5(2), 143–156.
- Johnston, W., & Uhl, C. (1976). The contributions of encoding effort and variability to the spacing effect on free recall. *Journal of Experimental Psychology: Human Learning and Memory*, 2(2), 153–160.
- Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fMRI studies. *NeuroImage*, 54(3), 2446–2461.
- Kim, M.-S., Kim, J.-J., & Kwon, J. n (2001). The effect of immediate and delayed word repetition on event-related potential in a continuous recognition task. *Cognitive Brain Research*, 11(3), 387–396.
- 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*, 108(14), 5903–5908.
- Kuhl, B., Shah, A., DuBrow, S., & Wagner, A. (2010). Resistance to forgetting associated with hippocampus-mediated reactivation during new learning. *Nature Neuroscience*, 13(4), 501–506.
- Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, 48(6), 609–621.
- Levy, D., Stark, C., & Squire, L. (2004). Intact conceptual priming in the absence of declarative memory. *Psychological Science*, 15(10), 680–686.
- Liu, Q., Dong, Q., Chen, C. S., & Xue, G. (2014). Neural processes during encoding support durable memory. *NeuroImage*, 88, 1–9.
- Lu, Y., Wang, C., Chen, C., & Xue, G. (2015). Spatiotemporal neural pattern similarity supports episodic memory. *Current Biology*, 25(6), 780–785.
- Madigan, S. (1969). Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior*, 8(6), 828–835.
- Mammarella, N., Russo, R., & Avons, S. (2002). Spacing effects in cued-memory tasks for unfamiliar faces and nonwords. *Memory & Cognition*, 30(8), 1238–1251.
- Manning, J. R., Polyn, S. M., Baltuch, G. H., Litt, B., & Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Sciences*, 108(31), 12893–12897.
- Marini, F., Marzi, T., & Viggiano, M. P. (2011). “Wanted!” The effects of reward on face recognition: electrophysiological correlates. *Cognitive, Affective, & Behavioral Neuroscience*, 11(4), 627–643.
- McKone, E. (1995). Short-term implicit memory for words and nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(5), 1108.
- McKone, E., & Dennis, C. (2000). Short-term implicit memory: Visual, auditory, and cross-modality priming. *Psychonomic Bulletin & Review*, 7(2), 341–346.

- Mei, L., Xue, G., Chen, C., Xue, F., Zhang, M., & Dong, Q. (2010). The “visual word form area” is involved in successful memory encoding of both words and faces. *NeuroImage*, 52(1), 371–378.
- Miller, J. F., Neufang, M., Solway, A., Brandt, A., Trippel, M., Mader, I., et al. (2013). Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science*, 342(6162), 1111–1114.
- Mumford, J. (2007). *A guide to calculating percent change with feaquery*. Unpublished Tech Report In http://mumford.bol.ucla.edu/perchange_guide.pdf.
- Murray, J. (1983). *Spacing phenomena in human memory: A study-phase retrieval interpretation*. Doctoral dissertation, University of California, Los Angeles. Available from Dissertation Abstracts International.
- Nahum, L., Gabriel, D., Spinelli, L., Momjian, S., Seeck, M., Michel, C. M., et al. (2011). Rapid consolidation and the human hippocampus: intracranial recordings confirm surface EEG. *Hippocampus*, 21(7), 689–693.
- Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage*, 25(3), 653–660.
- Nissen, M. J., Knopman, D. S., & Schacter, D. L. (1987). Neurochemical dissociation of memory systems. *Neurology*, 37(5), 789–789.
- Otten, L. J., Henson, R. N., & Rugg, M. D. (2001). Depth of processing effects on neural correlates of memory encoding relationship between findings from across-and within-task comparisons. *Brain*, 124(2), 399–412.
- Otten, L. J., Sveen, J., & Quayle, A. H. (2007). Distinct patterns of neural activity during memory formation of nonwords versus words. *Journal of Cognitive Neuroscience*, 19(11), 1776–1789.
- Padovani, T., Koenig, T., Brandeis, D., & Perrig, W. J. (2011). Different brain activities predict retrieval success during emotional and semantic encoding. *Journal of Cognitive Neuroscience*, 23(12), 4008–4021.
- Paller, K. A., Voss, J. L., & Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends in cognitive sciences*, 11(6), 243–250.
- Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science*, 310(5756), 1963–1966.
- Raaijmakers, J. (2003). Spacing and repetition effects in human memory: application of the SAM model. *Cognitive Science*, 27(3), 431–452.
- Race, E. A., Shanker, S., & Wagner, A. D. (2009). Neural priming in human frontal cortex: multiple forms of learning reduce demands on the prefrontal executive system. *Journal of Cognitive Neuroscience*, 21(9), 1766–1781.
- Reder, L. M., Park, H., & Kieffaber, P. D. (2009). Memory systems do not divide on consciousness: reinterpreting memory in terms of activation and binding. *Psychological bulletin*, 135(1), 23.
- Richardson-Klavehn, A., & Bjork, R. A. (1988). Primary versus secondary rehearsal in an imagined voice: differential effects on recognition memory and perceptual identification. *Bulletin of the Psychonomic Society*, 26(3), 187–190.
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in cognitive sciences*, 11(6), 251–257.
- Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, 392(6676), 595–598.
- Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends in cognitive sciences*, 7(7), 313–319.
- Russo, R., Mammarella, N., & Avons, S. (2002). Toward a unified account of spacing effects in explicit cued-memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(5), 819.
- Russo, R., Parkin, A. J., Taylor, S. R., & Wilks, J. (1998). Revising current two-process accounts of spacing effects in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(1), 161–172.
- Schott, B. H., Richardson-Klavehn, A., Henson, R. N., Becker, C., Heinze, H.-J., & Düzel, E. (2006). Neuroanatomical dissociation of encoding processes related to priming and explicit memory. *The Journal of Neuroscience*, 26(3), 792–800.
- Shimamura, A. P. (2011). Episodic retrieval and the cortical binding of relational activity. *Cognitive, Affective, & Behavioral Neuroscience*, 11(3), 277–291.
- Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of learning and memory*, 82(3), 171–177.
- Stark, S. M., Gordon, B., & Stark, C. E. (2008). Does the presence of priming hinder subsequent recognition or recall performance? *Memory*, 16(2), 157–173.
- Thios, S. J., & Dagostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior*, 15(5), 529–536.
- Toppino, T. C., Fearnow-Kenney, M. D., Kiepert, M. H., & Teremula, A. C. (2009). The spacing effect in intentional and incidental free recall by children and adults: limits on the automaticity hypothesis. *Memory & Cognition*, 37(3), 316–325.
- Turk-Browne, N. B., Yi, D.-J., & Chun, M. M. (2006). Linking implicit and explicit memory: common encoding factors and shared representations. *Neuron*, 49(6), 917–927.
- Uncapher, M. R., & Rugg, M. D. (2005). Effects of divided attention on fMRI correlates of memory encoding. *Journal of Cognitive Neuroscience*, 17(12), 1923–1935.
- Van Essen, D. C. (2005). A Population-Average, landmark- and surface-based (PALS) atlas of human cerebral cortex. *NeuroImage*, 28(3), 635–662.
- Van Strien, J. W., Hagenbeek, R. E., Stam, C. J., Rombouts, S. A., & Barkhof, F. (2005). Changes in brain electrical activity during extended continuous word recognition. *NeuroImage*, 26(3), 952–959.
- Van Strien, J. W., Verkoeijen, P. P., Van der Meer, N., & Franken, I. H. (2007). Electrophysiological correlates of word repetition spacing: ERP and induced band power old/new effects with massed and spaced repetitions. *International Journal of Psychophysiology*, 66(3), 205–214.
- Verkoeijen, P. P. J. L. (2005). *Explaining the spacing effect: Study-phase retrieval, contextual-variability, and priming accounts*. Erasmus University Rotterdam.
- Voss, J. L., & Paller, K. A. (2006). Fluent conceptual processing and explicit memory for faces are electrophysiologically distinct. *The Journal of Neuroscience*, 26(3), 926–933.
- Voss, J. L., & Paller, K. A. (2007). Neural correlates of conceptual implicit memory and their contamination of putative neural correlates of explicit memory. *Learning & Memory*, 14(4), 259–267.
- Voss, J. L., & Paller, K. A. (2009). Remembering and knowing: electrophysiological distinctions at encoding but not retrieval. *NeuroImage*, 46(1), 280–289.
- Voss, J. L., Schendan, H. E., & Paller, K. A. (2010). Finding meaning in novel geometric shapes influences electrophysiological correlates of repetition and dissociates perceptual and conceptual priming. *NeuroImage*, 49(3), 2879–2889.
- Wagner, A. D., Maril, A., & Schacter, D. L. (2000). Interactions between forms of memory: when priming hinders new episodic learning. *Journal of Cognitive Neuroscience*, 12(Suppl. 2), 52–60.
- Ward, E. J., Chun, M. M., & Kuhl, B. A. (2013). Repetition suppression and multi-voxel pattern similarity differentially track implicit and explicit visual memory. *The Journal of Neuroscience*, 33(37), 14749–14757.

- Woolrich, M. (2008). Robust group analysis using outlier inference. *NeuroImage*, 41(2), 286–301.
- Woolrich, M. W., Behrens, T. E., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, 21(4), 1732–1747.
- Xue, G., Dong, Q., Chen, C., Lu, Z., Mumford, J. A., & Poldrack, R. A. (2010a). Greater neural pattern similarity across repetitions is associated with better memory. *Science*, 330(6000), 97–101.
- Xue, G., Dong, Q., Chen, C., Lu, Z.-L., Mumford, J. A., & Poldrack, R. A. (2013). Complementary role of frontoparietal activity and cortical pattern similarity in successful episodic memory encoding. *Cerebral Cortex*, 23(7), 1562–1571.
- Xue, G., Mei, L., Chen, C., Lu, Z.-L., Poldrack, R., & Dong, Q. (2010b). Facilitating memory for novel characters by reducing neural repetition suppression in the left fusiform cortex. *PLoS One*, 5(10).
- Xue, G., Mei, L., Chen, C., Lu, Z.-L., Poldrack, R., & Dong, Q. (2011). Spaced learning enhances subsequent recognition memory by reducing neural repetition suppression. *Journal of Cognitive Neuroscience*, 23(7), 1624–1633.
- Yi, D.-J., & Chun, M. M. (2005). Attentional modulation of learning-related repetition attenuation effects in human parahippocampal cortex. *The Journal of Neuroscience*, 25(14), 3593–3600.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: a review of 30 Years of research. *Journal of Memory and Language*, 46(3), 441–517.