Research report

Cross-modal versus within-modal recall: Differences in behavioral and brain responses

Andrew J. Butler [a,b,*], Karin H. James [a,b,c]

[a] Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405, United States
[b] Program in Neural Science, Indiana University, Bloomington, IN 47405, United States
[c] Cognitive Science Program, Indiana University, Bloomington, IN 47405, United States

ABSTRACT

Although human experience is multisensory in nature, previous research has focused predominantly on memory for unisensory as opposed to multisensory information. In this work, we sought to investigate behavioral and neural differences between the cued recall of cross-modal audiovisual associations versus within-modal visual or auditory associations. Participants were presented with cue-target associations comprised of pairs of nonsense objects, pairs of nonsense sounds, objects paired with sounds, and sounds paired with objects. Subsequently, they were required to recall the modality of the target given the cue while behavioral accuracy, reaction time, and blood oxygenation level dependant (BOLD) activation were measured. Successful within-modal recall was associated with modality-specific reactivation in primary perceptual regions, and was more accurate than cross-modal retrieval. When auditory targets were correctly or incorrectly recalled using a cross-modal visual cue, there was re-activation in auditory association cortex, and recall of information from cross-modal associations activated the hippocampus to a greater degree than within-modal associations. Findings support theories that propose an overlap between regions active during perception and memory, and show that behavioral and neural differences exist between within- and cross-modal associations. Overall the current study highlights the importance of the role of multisensory information in memory.

1. Introduction

Humans have the ability to associate and later retrieve information within the same modality (within-modal associations) as well as across different modalities (cross-modal associations). However, there is little empirical work assessing the behavioral and neural impact of encoding multisensory information on subsequent memory processes. This should not be ignored because we have evolved neural systems to take advantage of the rich multisensory information encountered in our environment [1]. Neuroimaging studies reveal that widespread multisensory networks activate during multisensory perception [30], and that these systems may allow for behavioral enhancements during multisensory identification compared to unisensory identification [31]. Previous studies have shown that multisensory information modifies both behavioral and neuroimaging measures during subsequent unisensory presentations even after single-trail exposures [45,2]. More extensive training of novel cross-modal associations also alters brain regions

involved in audiovisual integration and congruency processing [3]. Therefore, just as it is important to consider multisensory processing during perception it is equally important to consider its role in memory processes. The current study aims to further investigate the role of unisensory regions and multisensory associative brain regions during cued-recall to better understand both the unisensory and multisensory nature of memory processes.

Although traditional approaches have maintained that perception and memory are separate systems [4], recent theory and research suggest that perceptual and memory processes are highly interdependent [6–8,32,5,37,43]. For example, Barsalou's [32] theory of perceptual symbols systems proposed that simulations in the modality specific perceptual regions of the brain are essential for memory. Fuster [37] stressed the importance of the reactivation of distributed hierarchal networks existing within and between modality specific perceptual and motor regions for memory processes. The earlier work of Damasio [5] has been influential in several of these theories. It too suggested that memory becomes represented in modality specific regions involved in perception and action, but also provided a possible neural mechanism by which this may occur. Specifically, in this theory, memory recall and recognition involve the reinstatement of activity in modality-specific perceptual and motor regions by means of higher order regions that

* Corresponding author at: Indiana University, 107 S Indiana Avenue, Bloomington, IN 47405, United States. Tel.: +1 812 856 7237.
E-mail address: Butler7@indiana.edu (A.J. Butler).

store patterns of perceptual or motor related activity. Accumulating evidence over the last two decades provides general support for important aspects of this theory suggesting that memory processes are intimately tied to perceptual encoding [6].

An important aspect of these theories is the proposal that regions related to processing information from specific perceptual modalities are also involved during later recall and recognition. This has been termed by some as the reactivation hypothesis of memory [11]. The reactivation of information represented in modality-specific regions may provide the specific content of memory retrieval [9]. Reactivation may also be related to the notion that memory relies on the simulation of perception [33], and the more general proposal that memory retrieval involves the reinstatement of cortical activity that occurred during encoding [8]. Single-unit recordings in monkeys have provided support for the importance of perceptual reactivation in visual memory [10]. Furthermore, research into conceptual deficits suggests that lesions of specific motor and sensory regions of the brain interfere with knowledge of certain concepts [5].

Neuroimaging studies have provided evidence for the occurrence of modality-specific reactivation during memory tasks (for a review see Ref. [49]). Using techniques such as PET or fMRI, studies have demonstrated reactivation occurring in both visual and auditory regions [41,11], visual regions alone [12], auditory regions alone [46], motor/somatosensory regions [13], and olfactory regions [14]. The modulation of these reactivation effects by various factors has also been investigated. For example, recognition associated with a 'remember' (as opposed to a 'know') response is associated with greater reactivation of sensory regions that are specific to associated contextual information encountered during encoding [15,22]. This greater activation of context related regions is associated with an increase in recognition accuracy [48], and recalling more information has been shown to increase the degree of neural reactivation [16].

The veracity of a memory also modulates the degree of perceptual reactivation. One study using an item and source recognition paradigm demonstrated that false alarms, indexing false memories, still show reactivation of encoding related areas specific to the false information being retrieved [25]. In this study false alarms, measuring false memories, led to reactivation effects in the same areas as veridical recognition as indexed by hit rate. These false reactivation effects were weaker, but still significant, when compared to veridical reactivation. Other studies have demonstrated activation differences in perceptual regions associated with true and false memories during both encoding and recognition. For example, activation in associative visual regions during encoding has been shown to be associated with both true and false memory formation, whereas primary visual regions were associated with true memory formation alone [17]. Furthermore, during recognition, associative visual regions showed activation during both true and false recognition of visual items, but primary visual regions showed activation only during true recognition [51].

Finally, it is important to consider the role of higher order multimodal associative regions in memory retrieval and perceptual reactivation. Some theories suggest that both within and cross-modal reactivation generally occur through pathways that involve the hippocampus [8]. Other theories suggest that cross-modal reactivation may rely to a greater extent on such multisensory processing regions whereas within-modal reactivation may rely more on regions that lie within modality-specific associative regions [5,6]. Furthermore, recent work has shown that the encoding of cross-modal associations shows greater activation of the hippocampus than the encoding of within-modal associations [39]. This difference may also be reflected during subsequent cued-recall.

In the current study, we investigated the effects of associating novel items that were presented in different modalities (visual

and auditory) on brain activation patterns during both encoding and cued recall. The current study included both auditory and visual cues that were paired with both auditory and visual targets. This design allowed us to test for differences between the recall of cross-modal versus within-modal information. We also compared correct to incorrect recall performance to test for false reactivation effects. Thus the design of the current study allows us to test the reactivation hypothesis in several ways while at the same time exploring the effects of retrieving cross-modal versus within-modal associations during both correct and incorrect recall. Based on previous findings we had several expectations. First, we expected that modality-specific reactivation in visual or auditory regions would occur during the correct recall of visual or auditory information, respectively. Secondly, we expected that non-primary perceptual regions might show false reactivation effects. In other words, the recall of incorrect information would lead to reactivation in the modality-specific visual or auditory regions associated with the incorrectly recalled visual or auditory target. Finally, we expected greater recruitment of regions crucial for relational binding, such the hippocampus, during the recall of cross-modal versus within-modal information.

2. Methods

2.1. Participants

Twelve individuals (8 females and 4 males) participated in the study (mean age = 23.4, SD = 3.4). All gave informed consent according to the guidelines of the Indiana University institutional review board. All participants were right handed, and reported normal or corrected-to-normal vision. Participants were compensated for their time.

2.2. Stimuli

Stimuli used were a total of 64 different visual images of novel 3D objects (see examples in Fig. 1) and a total of 64 different novel auditory sounds. Novel 3D objects were presented, in grayscale, on the center of the screen as 1 s static images. Novel auditory stimuli were presented binaurally for 1 s, at a sound level above the MRI scanner noise so that they could be perceived (as determined by pre-testing). Novel auditory sounds were auditory clips of machine sounds that were scrambled using a MATLAB program that randomly mixed portions of the original sound files so as to make them unrecognizable.

2.3. Procedure

After informed consent, MRI safety screening, and initial instructions participants entered the MRI scanner, and began the experiment. Over the entire course of the experiment a total of 11 experimental runs were presented. The first 2 runs consisted of visual and auditory presentations that served to functionally localize modality specific regions of the brain. As such, participants simply viewed the objects or listened to the sounds. Stimulus items were used that were different from those used during the following encoding and recall runs. Runs 3 through 8 consisted of the encoding of stimulus pair associations. These visual and auditory stimuli were
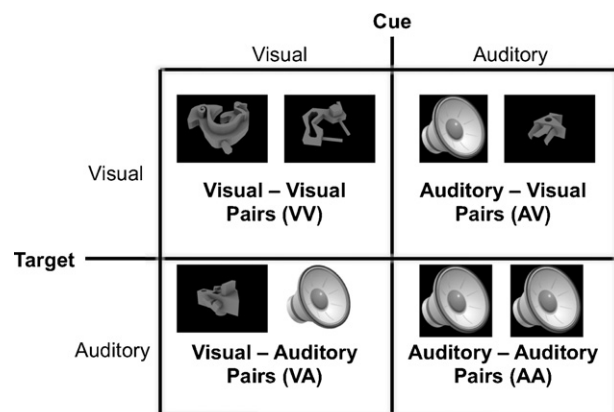

Fig. 1. Stimuli and design. Pairing of visual and auditory stimuli in experimental design.

paired in four different ways (see Fig. 1 for this aspect of the design). The four types of pairs included 16 pairs of visual objects (VV pairs), 16 pairs of sounds (AA pairs), 16 visual objects paired with sounds (VA), and 16 sounds paired with visual objects (AV). To control for the effect of specific items all individual stimuli were counterbalanced between position as a cue or target, and between pair condition type (VV, AA, VA, and AV). During encoding all participants were exposed to all of the pairs across all conditions 6 times. The task during the encoding runs was passive viewing and listening while attempting to learn the associations. Because the cue and the target were each 1 s and the target was presented immediately after the cue, the duration of a pair during encoding equaled 2 s. The ISI during the encoding runs was jittered at 2, 4, or 6 s to allow for rapid presentation. Runs 9 and 10 consisted of a slow event-related memory task in which participants were presented with cues alone and responded whether the cue was previously paired with an auditory or visual target. During these runs the first item in a pair from the encoding runs served as a cue to recall information about the second target item in a pair. Each audio or visual cue lasted for 1 s followed by another second of fixation. The ISI during these runs was set at 12 s. Participants used the index and middle finger of their right hand to make their response. The final run consisted of a high-resolution anatomical scan.

2.4. Imaging parameters

Imaging was performed using a 3-T Siemens Magnetom Trio whole body MRI system and a phased array eight channel head coil, located at the Indiana University Psychological and Brain Sciences department. All stimuli were back-displayed via a Mitsubishi XL30 projector onto a screen that was viewed through a mirror from the bore of the scanner. Stimuli were presented via MATLAB software via a Macintosh Macbook laptop.

The field of view was 22 cm × 22 cm × 9.9 cm, with an in plane resolution of 64 × 64 pixels and 33 slices per volume that were 3.4 mm thick. These parameters allowed us to collect data from the entire brain. The resulting voxel size was 3.4 mm × 3.4 mm × 3.4 mm. Images were acquired using an echo-planar technique (TE = 30 ms TR = 2000 ms, flip angle = 90°) for BOLD based imaging. High-resolution T1-weighted anatomical volumes were acquired using a 3D Turbo-flash acquisition.

2.5. fMRI data analysis procedures

Brain Voyager™ (Brain Innovation, Maastricht, Netherlands) was used to analyze the fMRI data. During preprocessing functional data underwent slice time correction, 3D motion correction, linear trend removal, and Gaussian spatial blurring (FWHM 6 mm). Individual functional volumes were co-registered to anatomical volumes with an intensity-matching, rigid-body transformation algorithm. Individual anatomical volumes were normalized to the stereotactic space of Talairach and Tournoux [55] using an eight-parameter affine transformation, with parameters selected by visual inspection of anatomical landmarks. Applying the same affine transformation to the co-registered functional volumes placed the functional data in a common brain space, allowing comparisons across participants. Voxel size of the normalized functional volumes was resampled at 3 mm³ using trilinear interpolation. It was this voxel size to which the cluster-size threshold was applied. Brain maps in figures are shown with the voxel size resampled at 1 mm³.

After preprocessing, the data from the localizer and recall runs, each data set were entered into their own general linear models (GLMs) using an assumed hemodynamic response function. Statistical parametric maps (SPMs) for the localizer runs were created for each of the participants' data. From the localizer runs functional regions of interest (ROIs) were defined in each individual, in visual and auditory regions using 'visual stimuli presentation > rest' and 'auditory stimuli presentation > rest' contrasts, respectively. The four ROIs defined for each participant included left auditory area, right auditory area, left visual area, and right visual area. We were specifically interested in the reactivation of regions that are involved in visual and auditory perception. In order for participants to learn the various and auditory associations during encoding they must perceive the stimuli. This should mean that the information is processed in primary visual and auditory cortex. Thus we assume these regions are active during encoding as well as during the perceptual localizer. The most statistically significant cluster of activation was chosen in each of these regions. All ROIs were significant below p < .05, Bonferroni corrected. Visual ROIs were chosen based on their proximity to primary visual regions (calcarine sulcus). Auditory ROIs were chosen based on their proximity to primary auditory regions (Heschl's gyrus). See Supplementary Table 1 for individual participants ROI cluster information. Time course data was taken from each participant's recall runs in each of these individually defined ROIs. These time courses were derived from the average across all voxels of a given ROI. Within-subject, repeated measures ANOVA's were performed using peak BOLD response from 4 to 10 s post-stimulus onset in each ROI, for each participant, as the dependant measure (see Supplementary Table 2 for an ANOVA table of these results).

Whole-brain statistical parametric map (SPM) analysis involving several contrasts of interest from the functional imaging runs were performed and are described in detail in the results section. We used the BrainVoyager Cluster-Level Statistical Threshold Estimator plugin to control for multiple tests. The plugin estimates the cluster-size threshold necessary to produce an effective alpha < .05, given a specific voxel-wise p-value, using Monte Carlo simulation. The statistical significance of clusters in a given contrast was first assessed using a random-effects model. Voxel-wise

significance was set at p < .005. The Cluster-Level Statistical Threshold Estimator plugin estimated a cluster-size threshold of six 3 mm³ voxels.

3. Results

3.1. Behavioral results

Accuracy across the four conditions was measured during the memory recall and is presented in Fig. 2. A repeated measures ANOVA of behavioral accuracy across these four conditions showed no main effect of cue ($F(1,11) = .94$, ns) or target ($F(1,11) = .07$, ns), but revealed a significant interaction of cue and target type ($F(1,11) = 7.89$, $p < .05$). Post hoc t-tests revealed that the within-modal conditions (VV and AA) had significantly greater accuracy than the cross-modal conditions (VA and AV). The VV condition had greater accuracy than the VA condition ($t(11) = 3.2$, $p < .01$). The AA condition had greater accuracy than the AV condition ($t(11) = 2.9$, $p < .01$). The VV condition also had greater accuracy that the AV condition ($t(11) = 2.2$, $p < .05$).

Reaction times across all conditions are shown in Fig. 2. A repeated measure ANOVA revealed that there was no effect of correct versus incorrect performance. There was a significant main effect of cue type ($F(1,11) = 13.897$, $p < .01$), but no main effect of target type ($F(1,11) = .004$, ns) or interaction ($F(1,11) = 3.51$, ns). Post hoc t-tests revealed that visual cues led to faster correct recall than auditory cues ($t(11) = 2.7$, $p < .01$).
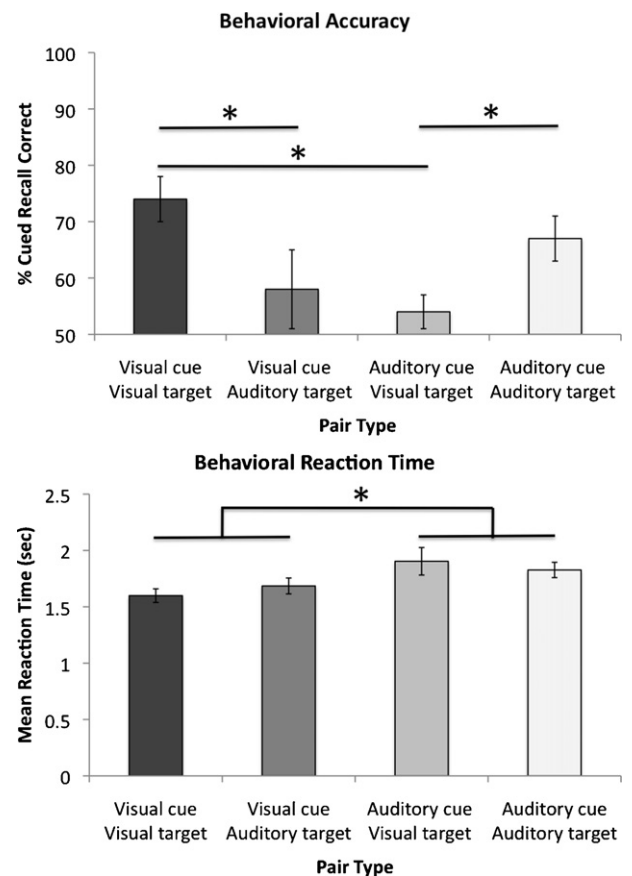

Fig. 2. Behavioral results. Accuracy and reaction time results for correct trials only. Within-modal cues were more accurately recalled than cross-modal cues. Visual cues were retrieved more quickly than auditory cues. In all graphs, error bars represent standard error of the mean. (*Statistically significant difference at p < .05 for all graphs.)

3.2. Functional ROI results

Reactivation effects related to the correct recall of targets were investigated in both auditory and visual regions. These auditory and visual perceptual ROIs revealed modality specific reactivation effects for within-modal associations. This type of reactivation occurs when a cue from a one modality is used to retrieve a target from the same modality. In visual regions this effect would be evidenced by greater activation of correctly retrieved VV targets relative to correctly retrieved VA targets. In auditory regions, this effect would be evidenced by greater activation of correctly retrieved AA targets relative to correctly retrieved AV targets. These results are detailed below. The ANOVAs for the ROI analysis were performed in two ways for each functional ROI. First, the ANOVAs for the ROI analyses were performed using data from only correct trials (i.e. there were four condition correct VV, correct VA, correct AV, and correct AA). In this ANOVA cue and target were the only factors. Secondly, we also wanted to perform ANOVAs that included both correct and incorrect responses modeled as separate conditions to ascertain whether there were effects or interactions related to the factor of accuracy. Accuracy was a within groups factor that included two levels – correct and incorrect. Accuracy therefore refers to correct or incorrect conditions and *not* to the percent correct on a scale from 0% to 100%.

3.2.1. Visual ROIs

In the left visual ROI there was greater activation during the correct recall of visual targets using visual cues (VV condition) than the correct recall of auditory targets using visual cues (VA condition) (see Fig. 3). A repeated measures ANOVA for correct performance in the left visual ROI demonstrated a significant main effect of both cue ($F(1,11) = 42.38$, $p < .001$) and target ($F(1,11) = 4.8$, $p < .05$), but no significant interaction ($F(1,11) = .07$, ns). A post hoc t-test revealed the response during VV was significantly greater than the response during VA ($t(11) = 2.7$, $p < .01$). However, the response during AV was not significantly different from the response during AA ($t(11) = 1.23$, $p = .244$). There was no significant main effect ($F(1,11) = 1.6$, $p = .233$) or interaction ($F(1,11) = .67$, $p = .430$) when accuracy was added as a factor.

A repeated measures ANOVA in the right visual ROI only revealed a significant main effect of cue ($F(1, 11) = 40.38$, $p < .001$) and no significant effect of target type ($F(1,11) = 1.14$, $p = .310$) or interaction ($F(1,11) = .367$, ns). There was no significant main effect ($F(1,11) = .063$, $p = .807$) or interaction ($F(1,11) = .395$, $p = .544$) when accuracy was added as a factor.

3.2.2. Auditory ROIs

In bilateral auditory regions there was greater activation during the correct recall of auditory targets using auditory cues (AA condition) than the correct recall of visual targets using auditory cues (AV condition) (see Fig. 4). A repeated measures ANOVA for correct performance in the left auditory ROI demonstrated a significant main effect of cue ($F(1,11) = 5.38$, $p = .03$), no significant main effect of target ($F(1,11) = 4.032$, ns), and a significant interaction of cue and target ($F(1,11) = 10.61$, $p < .01$). A post hoc t-test revealed a significant difference between AA correct and AV correct conditions ($t(13) = 3.0$, $p < .01$), such that AA correct had a greater peak response than AV correct in the left auditory ROI. However, the response during VA was not significantly different from the response during VV ($t(11) = .938$, $p = .368$.). There was no significant main effect ($F(1,11) = .070$, $p = .797$) or interaction ($F(1,11) = .004$, $p = .952$) when accuracy was added as a factor.
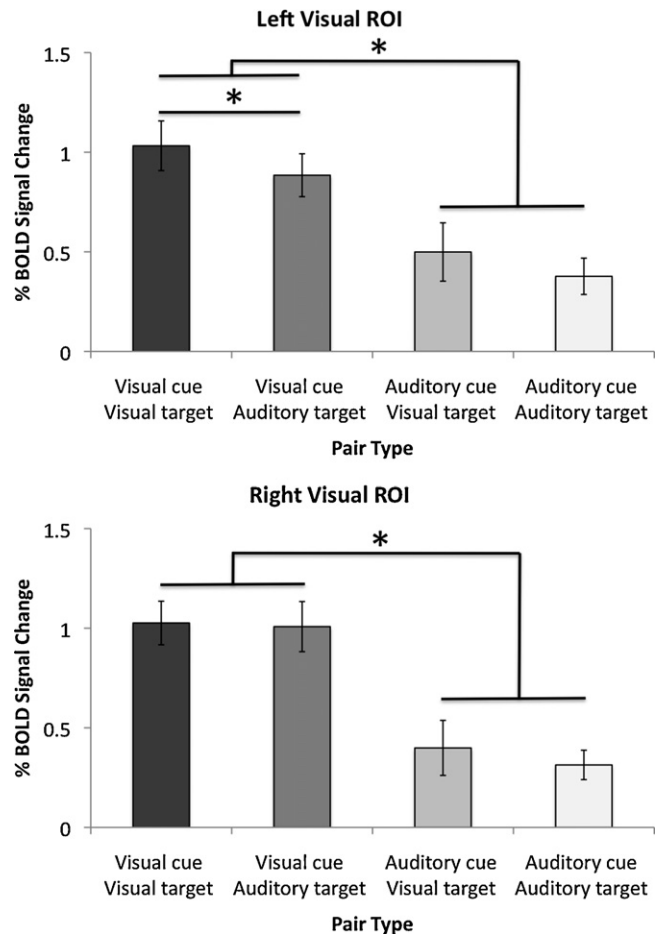

Fig. 3. Visual ROI results. Region-of-interest analyses in visual areas. The left visual ROI showed greater activation during the within-modal retrieval of visual information (VV) compared to the cross-modal retrieval of auditory information (VA). The right visual ROI did not show such a difference. Both left and right visual ROIs had greater activation during the presentation of visual compared to auditory cues.

A repeated measures ANOVA for correct performance in the right auditory ROI demonstrated a significant main effect of cue ($F(1,11) = 14.76$, $p < .005$) and target ($F(1,11) = 7.84$, $p < .02$) as well as a significant interaction of cue and target ($F(1,11) = 4.6$, $p < .05$). A post hoc t-test revealed a significant difference between AA correct and AV correct conditions ($t(13) = 3.0$, $p < .01$) such that AA correct also had a greater peak response than AV correct in the right-lateralized auditory ROI. However, the response during VA was not significantly different from the response during VV ($t(11) = 1.39$, $p = .193$). There was no significant main effect ($F(1,11) = 3.89$, $p = .077$) or interaction ($F(1,11) = .156$, $p = .701$) when accuracy was added as a factor.

3.2.3. Summary of ROI results

In summary the visual and auditory ROI analysis demonstrated that within-modal reactivation occurred in primary perceptual regions. For the visual modality this was evidenced by greater activation in visual ROIs when recalling visual targets using visual cues (VV) compared to recalling auditory targets using visual cues (VA). For the auditory modality this was evidenced by greater activation in auditory ROIs when recalling auditory targets using auditory cues (AA) compared to recalling visual targets using auditory cues (AV). Within-modal reactivation occurred in only left visual regions, but was bilateral in auditory regions. In these primary visual and auditory regions, there were no significant cross-modal reactivation effects (i.e. greater activation in visual or auditory regions when
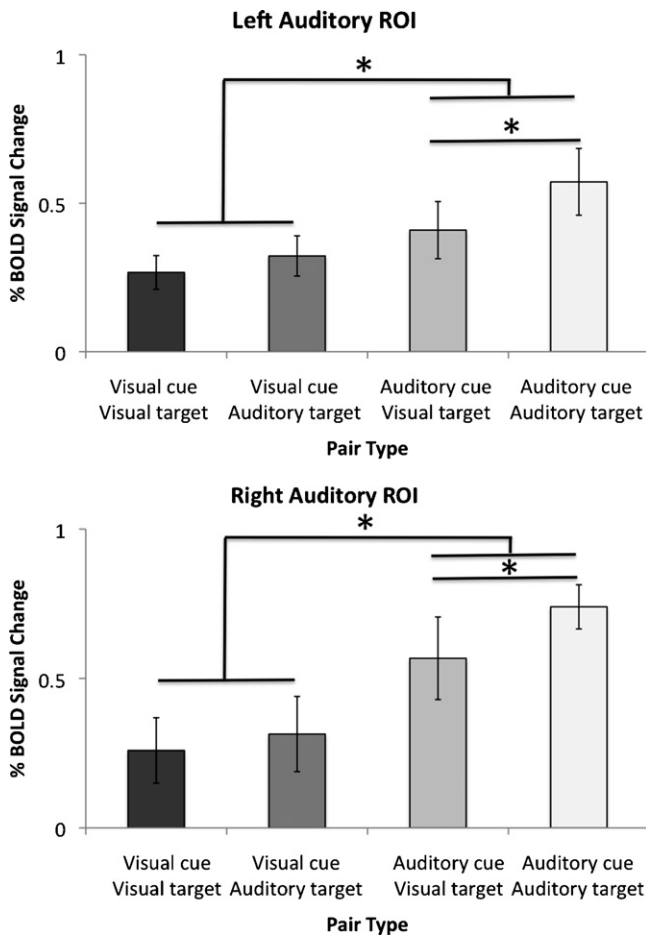
Fig. 4. Auditory ROI results. Region-of-interest analyses in auditory areas. Both left and right auditory ROIs showed significantly greater activation during the within-modal retrieval of auditory information (AA) than cross-modal retrieval (AV). Both left and right auditory ROIs had greater activation during the presentation of auditory compared to visual cues.

recalling visual or auditory targets using a cross-modal cue). Furthermore, there were no significant effects related to correct versus incorrect behavior when behavioral accuracy was added as a factor in the ROI ANOVAs.

3.3. Results of whole brain contrasts

Whole brain voxel-wise SPMs were created to further explore possible within- and cross-modal reactivation effects during correct performance. In regions where these reactivation effects occurred we compared correct versus incorrect performance. We performed these tests to assess the possibility of any false reactivation effects (i.e. reactivation based solely on the modality of the associated target regardless of the correctness of retrieval). See Supplementary Table 3 for detailed information on significant regions of activation in the following contrasts.

3.3.1. Activation related to target recall

We first performed several contrasts to test for activity related to target recall. A contrast of VV > VA revealed significant activation of a region in the left medial occipital lobe ($p < .005$, corrected) (see Fig. 5A). The inverse of this contrast (VA > VV) revealed significant activation in bilateral superior temporal gyrus (STG), the left superior temporal sulcus (STS), and multiple regions in the posterior and anterior cingulate cortex ($p < .005$, corrected) (see Fig. 5B). A contrast of AA > AV showed greater activation in a region in the left STG ($p < .005$, corrected) (see Fig. 5C). The inverse contrast (AV > AA) showed greater activation in the left caudate ($p < .005$, corrected).

A conjunction contrast was performed to test for regions that were overall more strongly active during both the with-modal and cross-modal recall of auditory targets (a conjunction of (VA > VV) and (AA > AV)). This conjunction contrast showed significantly greater activation in left lateralized anterior STG during the correct recall of auditory targets compared to visual targets ($p < .005$, corrected) (Fig. 6A). Note that this same region was also more active during the recall of auditory targets than visual targets in the opposite right hemisphere, but it did not survive the statistical threshold. Furthermore, the inverse contrast did not show any significant regions of activation. A repeated measures ANOVA including data from the incorrect conditions during cross-modal recall was compared in this region. This ANOVA revealed no main effect of target ($F(1,11) = 1.56$, ns) and correctness ($F(1,11) = .53$, ns), but a significant interaction between target and correctness ($F(1,11) = 6.8$, $p < .05$). Post hoc t-tests showed that VA targets resulted in greater activation than VV targets during correct recall ($t(11) = 1.9$, $p < .05$). However, VV showed greater activation than VA when targets were incorrectly recalled ($t(11) = 2.4$, $p < .01$) (see Fig. 6B). In this auditory region there was greater activation when participants were given
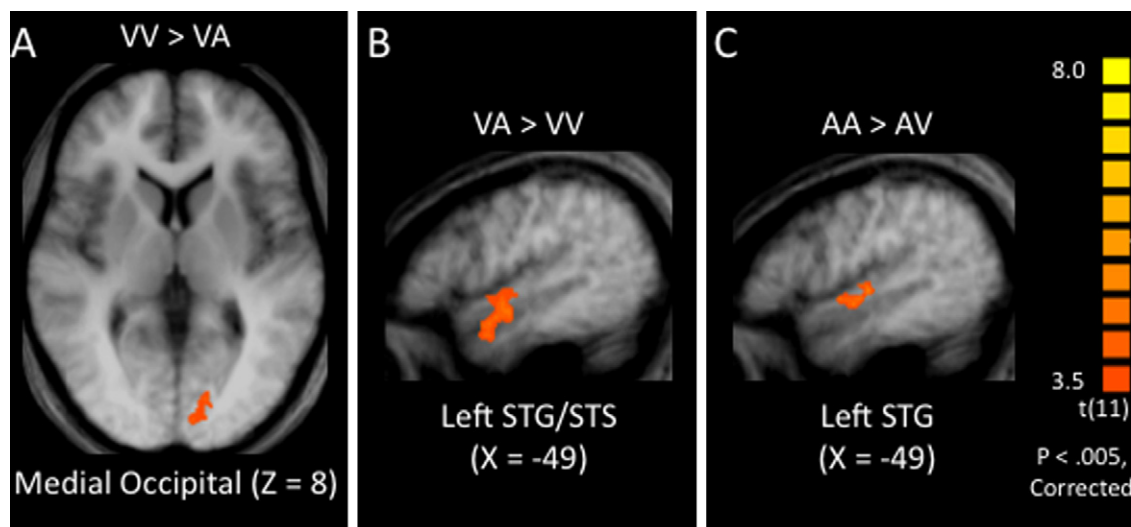

Fig. 5. Contrasts demonstrating reactivation effects. (A) Shows activation in a left medial occipital region as a result of the contrast of VV > VA ($p < .005$, corrected). (B) Shows activation in left STG/STG as a result of the contrast VA > AA ($p < .005$, corrected). (C) Shows activation in left STG as a result of the contrast AA > AV ($p < .005$, corrected).
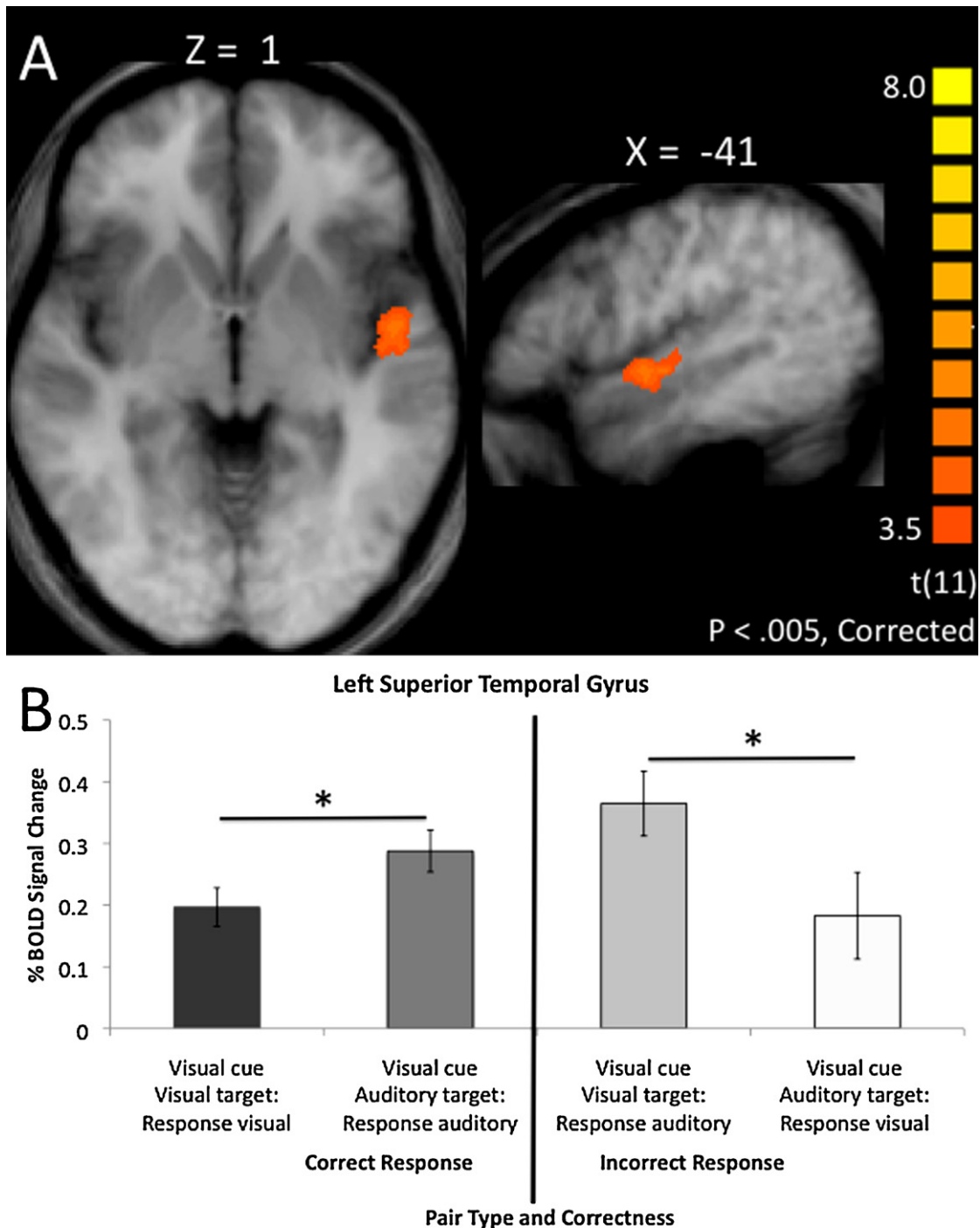
Fig. 6. Left STG reactivation and effect of incorrect performance. (A) Shows the results of the whole-brain conjunction contrast that was performed to test for regions that were overall more strongly active during both the with-modal and cross-modal recall of auditory targets (a conjunction of the contrasts (VA > VV) and (AA > AV)). The left superior temporal gyrus (STG) shows greater activation during the retrieval of auditory cues than visual cues ($p < .005$, corrected). The (B) (inset) shows the results for a comparison of correct versus incorrect retrieval using visual cues within this left STG region. This comparison shows that retrieving auditory target information led to greater activation than retrieving visual information, *regardless* of whether or not the retrieval was correct.

a visual cue and responded that the target was auditory regardless of whether the response was correct or not. In other words left lateralized STG activity increased whenever participants believed the target was presented through audition, independent of whether this was true.

3.3.2. Correct versus incorrect recall

We also directly compared the effects of correct compared to incorrect recall choices. A contrast of VV correct > VV incor-

rect showed greater activation of a region in the medial occipital lobe ($p < .005$, corrected) (see Fig. 7A). A contrast of VA correct > VA incorrect showed greater activation of bilateral regions in the STG as well as activation in the right posterior cingulate ($p < .005$, corrected) (see Fig. 7B). Similar contrasts using the AA and VV conditions comparing correct and incorrect performance did not show significant regions of activation. Also none of the inverse versions of these contrasts showed significant activation.
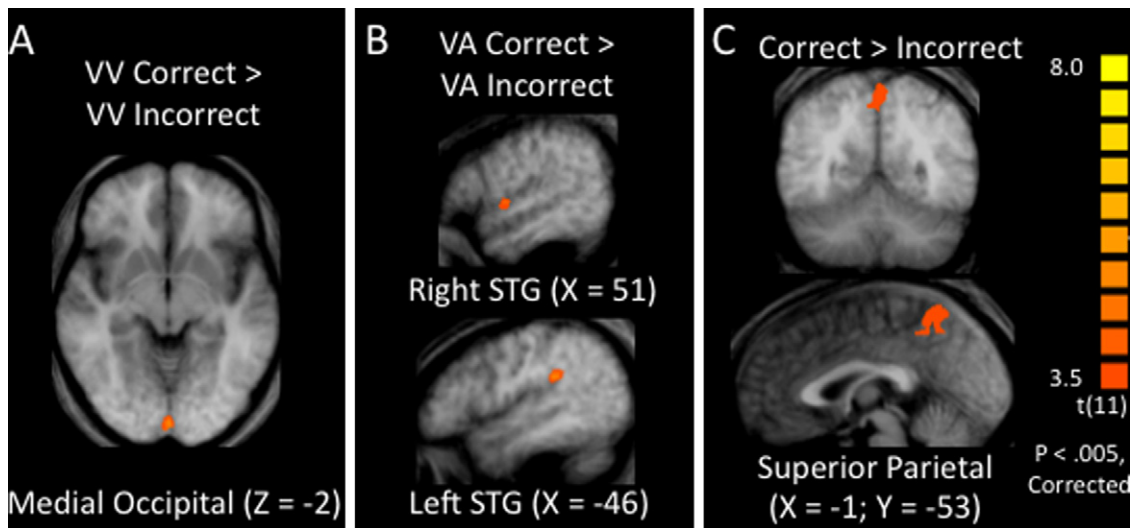
Fig. 7. Contrast of correct versus incorrect performance. (A) Shows activation in a medial occipital region as a result of the contrast of VV correct > VA incorrect ($p < .005$, corrected). (B) Shows activation in a bilateral STG regions as a result of the contrast of VA correct > VA incorrect ($p < .005$, corrected). (C) Shows the results of conjunction contrast (a conjunction of the following four contrasts: (VV correct > VV incorrect), (VA correct > VA incorrect), (AV correct > AV incorrect), and (AA correct > AA incorrect)) performed to compare correct to incorrect responses across all four types of pairs overall. This conjunction contrast revealed that activation a region in midline superior parietal cortex was stronger during correct recall versus incorrect recall ($p < .005$, corrected).

A conjunction contrast (a conjunction of the following four contrasts: (VV correct > VV incorrect), (VA correct > VA incorrect), (AV correct > AV incorrect), and (AA correct > AA incorrect)) was performed to compare correct to incorrect responses across all four types of pairs overall. The result of this contrast revealed that activation a region in midline superior parietal cortex was stronger during correct recall versus incorrect recall ($p < .005$, corrected) (see

Fig. 7C). The inverse contrast did not show any significant regions of activation.

3.3.3. Within- versus cross-modal recall

The final contrast of interest involved comparing within-modal to cross-modal recall conditions (Fig. 8). Fig. 8A shows conjunction contrast of correct cross-modal recall versus correct
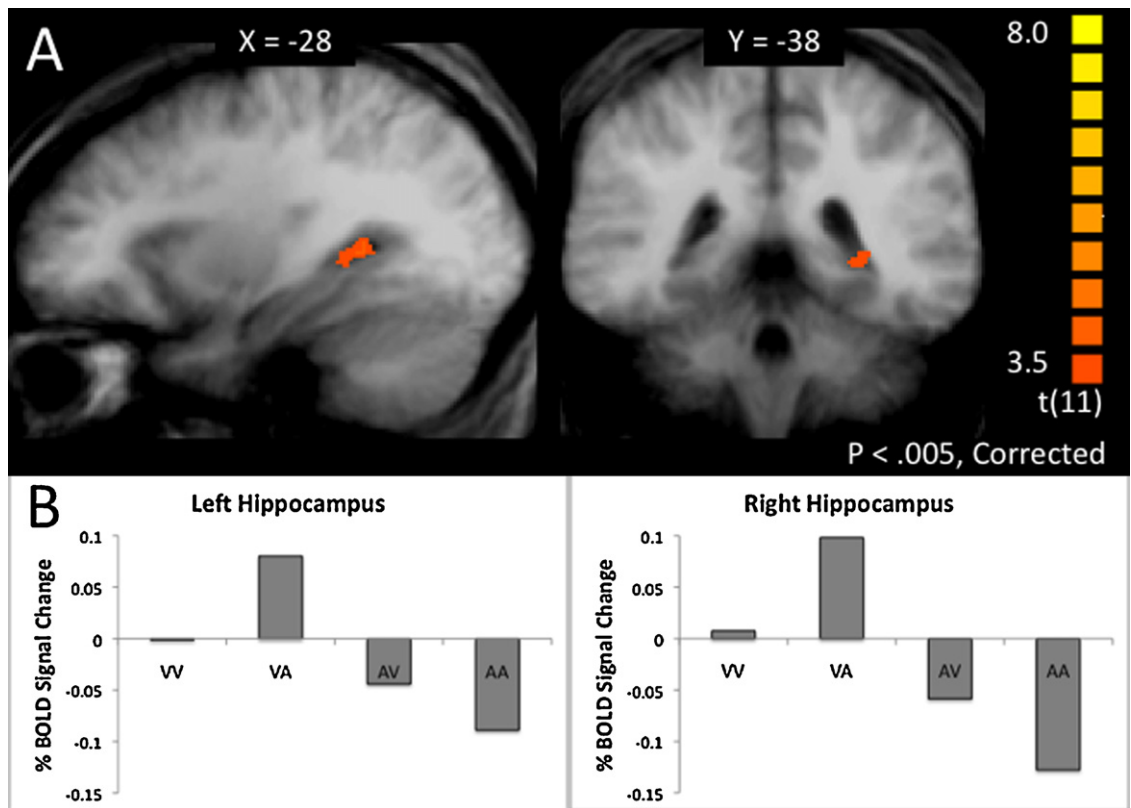

Fig. 8. Cross-modal versus within-modal activations. (A) Shows activation in right posterior hippocampus as a result of a conjunction contrast of correct cross-modal recall versus correct within-modal recall (the conjunction of (VA > VV) and (AV > AA)) revealed greater activation ($p < .005$, corrected) in the posterior left hippocampus (bilateral at lower thresholds) during cross-modal relative to within-modal recall. (B) Contains graphs showing the % BOLD signal change across correctly recalled conditions included for demonstration purposes.

within-modal recall (contrast: the conjunction of (VA > VV) and (AV > AA)) revealed greater activation (*p* < .005, corrected) in the posterior left hippocampus (bilateral at lower thresholds) during cross-modal relative to within-modal recall. As can be seen from Fig. 8B, in the % BOLD signal change plot of activation in both the left and right hippocampus, the cue modality modulated the interaction. The VA condition showed greater activation than the VV condition, and the AV condition showed greater activation than the AA condition. However, when the cues were visual the activation was greater than when they were auditory. The inverse contrast did not show significant regions of activation.

4. Discussion

4.1. Behavioral enhancement and modality-specific reactivation during within-modal retrieval

In the current study within-modal recall of visual or auditory targets led to significantly greater behavioral accuracy compared to the cross-modal recall of visual or auditory targets (see Fig. 2). This is the first study to test for and demonstrate such an effect. Previous work has demonstrated an increase in recognition accuracy and priming for pictures and sounds studied within a single modality [40]. The current study extends this finding of enhanced within-modal recognition to a cued recall task of modality specific information using novel auditory and visual stimuli. This finding is also generally in line with the theory of transfer-appropriate processing that suggests encoding and retrieval interact such that a greater overlap between processes occurring at retrieval and encoding enhances memory accuracy [44]. This accuracy advantage has been demonstrated empirically, and is associated with a greater overlap between encoding and retrieval processes in brain activation [18].

Based on the current fMRI results greater activation occurred in primary perceptual regions when participants correctly retrieved within-modal targets given within-modal cues (see Figs. 3 and 4). Because concurrent perception was controlled, this finding suggests that a memory related reactivation occurred in these perceptual regions. Specifically, using visual cues to correctly recall visual information was associated with reactivation of early left-lateralized occipital cortex, using auditory cues to correctly recall auditory information was associated with reactivation in bilateral posterior STG. These results support theories that maintain that memory retrieval requires the reactivation of perceptual regions engaged at encoding [5,8]. Importantly, similar effects in primary perceptual regions did not occur for cross-modal associations. However, as discussed in the next section, cross-modal reactivation effects were found in the anterior STG and STS during cross-modal auditory retrieval. The finding of within-modal reactivation in modality-specific visual and auditory regions also supports previous findings in studies investigating memory reactivation (e.g. [11]). Unlike this previous work, the current study used both novel audio and novel visual stimuli as both cues and targets in a task requiring the retrieval of target modality information. It is also unique in that it demonstrates within-modal auditory reactivation.

The reason why the visual reactivation effects were left lateralized in the current study is not readily apparent. However, similar left-lateralized effects are not without precedence. The clearest example is Wheeler et al. [11] who too found that while auditory reactivation effects were bilateral visual reactivation effects were left lateralized. In addition, several other studies have found left-lateralized activation of visual regions during the retrieval of visual information from memory (e.g. [19–22]; Slotnick et al., 2003). Due to the strong but not complete overlap between visual imagery and memory processes in visual cortex [52] it is important to con-

sider some relevant work into the nature of mental imagery as well. Specifically, previous lesion [36,23] and neuroimaging work [34] have suggested that left-lateralized visual regions may be specialized for mental image generation. Grossi and Modafferi [23] in particular propose that there is a division of labor between the left and right visual related hemispheres in which the left is necessary for generating visual images but the right is involved in the exploration visual images. In light of this it is possible that recall in this study only required visual image generation and not the exploration of visual images.

Crucially, the current results suggest that individuals may differentially encode and/or recall within-modal associations relative to cross-modal associations. The behavioral results, along with the within-modal reactivation of primary visual and auditory regions, suggest that the recall of within-modal associations may be more accurate because within-modal cues activate similar modality-specific regions in which the to-be-recalled target information is represented. Therefore, the current study provides one neural mechanism that would explain why an increased convergence between processes occurring at retrieval and encoding enhance memory accuracy.

4.2. Cross-modal auditory reactivation and effects of incorrect performance

Results from the whole-brain SPM analysis showed that activation of the left anterior STG occurred during both the cross-modal and within-modal retrieval of auditory target information (see Fig. 6A). Cross-modal auditory reactivation occurred in anterior superior temporal sulcus (STS) as well (see Fig. 5B). There are several proposed mechanisms by which cross-modal activation may occur during perceptual and memory tasks. One possibility is that modality-specific areas may be linked through medial temporal regions such as the hippocampus and perihinal cortex [53]. As discussed in detail later, the hippocampus was more greatly active during cross-modal retrieval suggesting that it may play a role in cross-modal retrieval and reactivation in the context of the present task. In addition, associative regions in the cortex such as the multisensory STS, among others, may play a role in linking modality specific activation during learning and memory [38]. In the current study the STS showed reactivation during cross-modal but not within-modal reactivation of auditory information (see Fig. 5B and C). Finally, direct feed-forward connections between modality-specific regions have been found to be involved in some cases of cross-modal reactivation [24]. Supporting this, recent work has shown that cross-modal repetition effects occur in earlier auditory and visual processing regions than has been suspected (Doehrmann et al., 2010).

In the current study we also looked at the effects of correct versus incorrect retrieval in regions that demonstrated reactivation effects. In the same left anterior STG region that demonstrated within- and cross-modal reactivation of target information what could be considered a 'false' reactivation effect was found (see Fig. 6B). When participants correctly or incorrectly recalled auditory targets there was greater activation in this region than when they recalled visual targets either correctly or incorrectly. Thus, at least in this region, reactivation seems to depend upon the type of target retrieved regardless of whether it was the correct one or not. This is in line with previous work that has shown that false alarms during recognition can be associated with reactivation effects in the same areas as veridical recognition [25]. The current study is the first to show this phenomenon can occur during cued retrieval of modality-specific auditory information. It should also be stressed that, as in previous work, this effect was demonstrated in more associative perceptual regions as opposed to primary perceptual

regions. Previous work has demonstrated that during encoding associative visual regions were recruited during true and false memory formation, and only primary visual regions were associated with true memory formation [17]. This finding also supports studies that have demonstrated that primary visual regions exhibit activations that distinguish between true and false visual recognition, but more associative visual regions do not [51]. One reason this effect was only significant in auditory regions, and not visual regions, may be due to the similarity of the auditory stimuli. Some theories have proposed that false memories are increased by the similarity between items (e.g. [26]). Experimental paradigms such as the Deese-Roediger-McDermott-Paradigm (DRM) [27] support this proposal, showing that false recognition increases when distractors are similar to studied words.

It is worth noting that when comparing correct versus incorrect performance across all conditions there was greater superior parietal activation (see Fig. 7C). Activations in both the inferior and superior parietal lobe have been repeatedly found in previous neuroimaging studies of memory when old, compared to new, items are correctly recalled or recognized [54].

4.3. Greater hippocampal activation during the retrieval of cross-modal target information

In the current study the posterior left hippocampus was more greatly active during the correct recall of cross-modal associations compared to the recall of within-modal associations (see Fig. 8A). This is the first evidence that the hippocampus is engaged to a greater degree during the recall of cross-modal information compared to within-modal information. The results also suggest that the particular cue and target pairings impacted the activation of the hippocampus in different ways. Specifically, while the pairs with visual cues had a relatively greater degree of activation than those with auditory cues, when the cues were the same the cross-modal pairing showed greater activation than the within-modal pairing (see Fig. 8B). Based on a large amount of empirical work to date the hippocampus is thought to play a critical role in relational memory [28]. Therefore, the increased activation in the hippocampus during the retrieval of cross-modal associations suggests an increased demand for relational memory for this type of association. This finding converges with work showing that the encoding of cross-modal associations is associated with greater activation of the hippocampus than the encoding within-modal associations [39]. Cross-modal retrieval may depend more on the hippocampus, in which multisensory information converges and associative links between different modalities can allow for cross-modal retrieval through the reactivation of cross-modal perceptual regions (e.g. anterior STG in this study). Finally, the fact that hippocampal activation was found in posterior regions is in line with previous work demonstrating that encoding processes tend to activate anterior regions of the hippocampus while recall processes tend to activate more posterior regions of the hippocampus [42,29,47].

5. Conclusions

The current study suggests that multisensory information impacts not only perception [1,30,31], but memory processes as well. Furthermore, findings from the current study support the reactivation hypothesis of memory. Within-modal reactivation occurred in both primary auditory and visual regions while cross-modal reactivation occurred only in auditory association areas. In addition, the current study suggests that, at least during cross-modal auditory retrieval, reactivation associated with a falsely retrieved target can occur in modality-specific associative regions. Furthermore, the retrieval of cross-modal associations

may increase the degree of hippocampal involvement compared to within-modal retrieval. Overall the current study demonstrates the behavioral and neural differences between within- and cross-modal associations, and highlights the importance of considering the role of multisensory information in memory.

Acknowledgments

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbr.2011.06.017.

References

[1] Meredith MA, Stein BE. Interactions among converging sensory inputs in the superior colliculus. Science 1983;221:389–91.
[2] Murray MM, Sperdin HF. Single-trial multisensory learning and memory retrieval. In: Naumer MJ, Kaiser J, editors. Multisensory object perception in the primate brain. LLC: Springer Science+Business Media; 2010.
[3] Naumer MJ, Doehrmann O, Muller NG, Muckli L, Kaiser J, Hein G. Cortical plasticity of audio-visual object representations. Cerebral Cortex 2009;19:1641–55.
[4] Atkinson RC, Shiffrin RM. The control of short-term memory. Scientific American 1971;225:82–90.
[5] Damasio AR. Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. Cognition 1989;33:37–43.
[6] Meyer K, Damasio A. Convergence and divergence in a neural architecture for recognition and memory. Trends in Neurosciences 2009;32:376–82.
[7] Rubin DC. The basic-systems model of episodic memory. Perspectives on Psychological Science 2006;1:277–311.
[8] Rugg MD, Johnson JD, Park H, Uncapher MR. Encoding-retrieval overlap in human episodic memory: a functional neuroimaging perspective. Progress in Brain Research 2008;169:339–52.
[9] Buckner RL, Wheeler ME. The cognitive neuroscience of remembering. Nature Reviews 2001;2:604–15.
[10] Miyashita Y, Hayashi T. Neural representation of visual objects: encoding and top-down activation. Current Opinion in Neurobiology 2000;10:187–94.
[11] Wheeler ME, Peterson SE, Buckner RL. Memory's echo: vivid remembering reactivates sensory-specific cortex. Proceedings of the National Academy of Sciences of the United States of America 2000;97:11125–9.
[12] Wheeler ME, Shulman GL, Buckner RL, Miezin FM, Velanova K, Peterson SE. Evidence for separate perceptual reactivation and search processes during remembering. Cerebral Cortex 2006;16:949–59.
[13] Nyberg L, Petersson KM, Nilsson LG, Sandblom J, Aberg C, Ingvar M. Reactivation of motor brain areas during explicit memory for actions. NeuroImage 2001;14:521–8.
[14] Gottfried JA, Smith APR, Rugg MD, Dolan RJ. Remembrance of odors past: human olfactory cortex in cross-modal recognition memory. Neuron 2004;42:687–95.
[15] Wheeler ME, Buckner RL. Functional-anatomic correlates of remembering and knowing. NeuroImage 2004;21:1337–49.
[16] Vilberg KL, Rugg MD. Dissociation of the neural correlates of recognition memory according to familiarity, recollection, and amount of recollected information. Neuropsychologia 2007;45:2216–25.
[17] Kim H, Cabeza R. Trusting our memories: dissociating the neural correlates of confidence in veridical versus illusory memories. The Journal of Neuroscience 2007;27:12190–7.
[18] Park H, Rugg MD. The relationship between study processing and the effects of cue congruency at retrieval: fMRI support for transfer appropriate processing. Cerebral Cortex 2008;18:868–75.
[19] Johnson JD, Rugg MD. Recollection and the reinstatement of encoding-related cortical activity. Cerebral Cortex 2007;17:2507–15.
[20] Kensinger EA, Schacter DL. Remembering the specific visual details of presented objects: neuroimaging evidence for effects of emotion. Neuropsychologia 2007;45:2951–62.
[21] Vaidya CJ, Zhao M, Desmond JE, Gabrieli JDE. Evidence for cortical encoding specificity in episodic memory: memory-induced re-activation of picture processing areas. Neuropsychologia 2002;40:2136–43.
[22] Woodruff CC, Johnson JD, Uncapher MR, Rugg MD. Content-specificity of the neural correlates of recollection. Neuropsychologia 2005;43:1022–32.

[23] Grossi D, Modafferi A. Different roles of the cerebral hemispheres in mental imagery: the "o'Clock Test" in two clinical cases. Brain and Cognition 1989;10:18–27.

[24] Foxe JJ, Schroeder CE. The case for feedfoward multisensory convergence during early cortical processing. NeuroReport 2005;16:419–23.

[25] Kahn I, Davachi L, Wagner AD. Functional-neuroanatomic correlates of recollection: implications for models of recognition memory. The Journal of Neuroscience 2004;24:4172–80.

[26] Reyna VF, Brainerd CJ. Fuzzy-trace theory: an interim synthesis. Learning and Individual Differences 1995;7:1–75.

[27] Roediger HL, McDermott KB. Creating false memories: remembering words not presented in lists. Journal of Experimental Psychology: Learning, Memory, and Cognition 1995;21:803–14.

[28] Konkel A, Cohen NJ. Relational memory and the hippocampus: representations and methods. Frontiers in Neuroscience 2009;3:166–74.

[29] Schacter DL, Wagner AD. Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. Hippocampus 1999;9:7–24.

[30] Amedi A, von Kriegstein K, Atteveldt NM, Beauchamp MS, Naumer MJ. Functional imaging of human crossmodal identification and object recognition. Experimental Brain Research 2005;166:559–71.

[31] Bolognini N, Frassinetti F, Serino A, Ladavas E. Acoustical vision" of below threshold stimuli: interaction among spatially converging audiovisual inputs. Experimental Brain Research 2005;160:273–82.

[32] Barsalou LW. Perceptual symbol systems. Behavioral and Brain Sciences 1999;22:577–609.

[33] Barsalou LW. Grounded cognition. Annual Review of Psychology 2008;59:617–45.

[34] D'Esposito M, Detre JA, Aguirre GK, Stallcup M, Alsop DC, Tippet LJ, et al. A function MRI study of mental image generation. Neuropsychologia 1997;35:725–30.

[36] Farah MJ, Levine DN. A case study of mental imagery deficit. Brain and Cognition 1988;8:147–64.

[37] Fuster JM. Cortex and memory. The Journal of Cognitive Neuroscience 2010;21:451–9.

[38] Gonzalo D, Shallice T, Dolan R. Time-dependent changes in learning audiovisual associations: a single-trial fMRI study. NeuroImage 2000;11:243–55.

[39] Gottlieb LJ, Uncapher MR, Rugg MD. Dissociation of the neural correlates of visual and auditory contextual encoding. Neuropsychologia 2010;48:137–44.

[40] Greene AJ, Easton RD, LaShell LSR. Visual-auditory events: cross-modal perceptual priming and recognition memory. Consciousness and Cognition 2001;10:425–35.

[41] Hornberger M, Rugg MD, Henson RNA. fMRI correlates of retrieval orientation. Neuropsychologia 2006;44:1425–36.

[42] Lepage M, Habib R, Tulving E. Hippocampal PET activations of memory encoding and retrieval: The HIPER model. Hippocampus 1998;8:313–22.

[43] Martin A. The representation of objects concepts in the brain. Annual Review of Psychology 2007;58:25–45.

[44] Morris CD, Bransford JD, Franks JJ. Levels of processing versus transfer appropriate processing. Journal of Verbal Learning and Verbal Behavior 1977;16:519–33.

[45] Murray MM, Foxe JJ, Wylie GR. The brain uses single-trial multisensory memories to discriminate without awareness. NeuroImage 2005;27:473–8.

[46] Nyberg L, Habib R, McIntosh AR, Tulving E. Reactivation of encoding-related brain activity during memory retrieval. Proceedings of the National Academy of Sciences of the United States of America 2000;9:11120–4.

[47] Prince SE, Daselaar SM, Cabeza R. Neural correlates of relational memory: successful encoding and retrieval of semantic and perceptual associations. The Journal of Neuroscience 2005;25:1203–10.

[48] Skinner EI, Grady CL, Fernandes MA. Reactivation of context-specific brain regions during retrieval. Neuropsychologia 2010;48:156–64.

[49] Slotnick SD. Visual memory and visual perception recruit common neural substrates. Behavioral and Cognitive Neuroscience Reviews 2004;3:207–21.

[51] Slotnick SD, Schacter DL. The nature of memory related activity in early visual areas. Neuropsychologia 2006;44:2874–86.

[52] Slotnick SD, Thompson WL, Kosslyn SM. Visual memory and visual mental imagery recruit common control and sensory regions of the brain. Cognitive Neuroscience, in press, doi:10.1080/17588928.2011.578210.

[53] Taylor KI, Moss HE, Stamatakis EA, Tyler LK. Binding crossmodal object features in perirhinal cortex. Proceedings of the National Academy of Sciences of the United States of America 2006;103:8239–44.

[54] Wagner AD, Shannon BJ, Kahn I, Buckner RL. Parietal lobe contributions to episodic memory retrieval. Trends in Cognitive Sciences 2005;9:445–53.

[55] Talairach J, Tournoux P. A Co-Planar Stereotactic Atlas of the Human Brain: 3-dimensional proportional system: An approach to cerebral mapping (M. Rayport, Trans.). New York: Thieme; 1988.