

Active Learning of Novel Sound-producing Objects: Motor Reactivation and Enhancement of Visuo-motor Connectivity

Andrew J. Butler^{1,2} and Karin Harman James¹

Abstract

■ Our experience with the world commonly involves physical interaction with objects enabling us to learn associations between multisensory information perceived during an event and our actions that create an event. The interplay among active interactions during learning and multisensory integration of object properties is not well understood. To better understand how action might enhance multisensory associative recognition, we investigated the interplay among motor and perceptual systems after active learning. Fifteen participants were included in an fMRI study during which they learned visuo-auditory-motor associations between novel objects and the sounds they produce, either through self-generated actions on the objects (active learning) or by observing an experimenter produce the actions (passive learning). Immediately after learning, behavioral and BOLD fMRI measures were collected while perceiving the objects used during unisensory and multisensory training in asso-

ciative perception and recognition tasks. Active learning was faster and led to more accurate recognition of audiovisual associations than passive learning. Functional ROI analyses showed that in motor, somatosensory, and cerebellar regions there was greater activation during both the perception and recognition of actively learned associations. Finally, functional connectivity between visual- and motor-related processing regions was enhanced during the presentation of actively learned audiovisual associations. Overall, the results of the current study clarify and extend our own previous work [Butler, A. J., James, T. W., & Harman James, K. Enhanced multisensory integration and motor reactivation after active motor learning of audiovisual associations. *Journal of Cognitive Neuroscience*, 23, 3515–3528, 2011] by providing several novel findings and highlighting the task-based nature of motor reactivation and retrieval after active learning. ■

INTRODUCTION

For typically developing individuals, interactions with the environment involve encoding information from multiple sensory inputs that are produced by self-generated actions—be they eye movements or effector movements. These interactions allow us to establish numerous associations, within and between various senses, connecting perceived objects and perceived outcomes in the context of goal-directed actions. This complex interplay of action and perception necessarily leads to interactions among neural systems for the production of efficient human behavior. Although coding multisensory information during a learning event facilitates subsequent recognition (Murray & Sperdin, 2010; Bolognini, Frassinetti, Serino, & Ladavas, 2005; Murray, Foxe, & Wylie, 2005) and acting on objects during learning events facilitates subsequent unisensory item recognition (Sasaoka, Asakura, & Kawahara, 2010; James et al., 2002; James, Humphrey, & Goodale, 2001; Harman, Humphrey, & Goodale, 1999), less is known about whether action during learning also facilitates multisensory associative recognition.

Neuroimaging studies have demonstrated that during perception and memory tasks, the reactivation of encoding-related regions occurs for the visual modality (Vilberg & Rugg, 2007; Hornberger, Rugg, & Henson, 2006; Slotnick & Schacter, 2006; Wheeler et al., 2006; Kahn, Davachi, & Wagner, 2004; Vaidya, Zhao, Desmond, & Gabrieli, 2002; Wheeler, Peterson, & Buckner, 2000), the auditory modality (Hornberger et al., 2006; James & Gauthier, 2003; Nyberg, Habib, McIntosh, & Tulving, 2000; Wheeler et al., 2000), the olfactory modality (Gottfried, Smith, Rugg, & Dolan, 2004), and with motor-related regions (e.g., Nyberg et al., 2001). The pervasive finding that active interaction with objects facilitates learning and results in the activation of motor systems upon subsequent perception supports theories that propose that representations of experience and knowledge include embodied information, derived during encoding, that is stored and subsequently reactivated in the brain's sensory and motor systems (e.g., Fuster, 2009; Barsalou, 1999; Damasio, 1989).

Although most studies to date have focused on the impact of active learning on subsequent recognition of stimuli processed using one modality, in everyday life, multiple modalities are used during perception and recognition of stimuli. Because active learning, involving self-generated

¹Indiana University, ²Butler University

actions on the environment, occurs in the context of multisensory perception during most everyday experiences, it may facilitate learning relative to passively (watching without overt action) gained experience. Furthermore, active learning may serve to facilitate associations being formed among sensory modalities (multisensory associative recognition). Recently, Butler, James, and Harman James (2011) addressed the impact of active learning on the subsequent perception and recognition of audiovisual (multisensory) associations. Participants in this study learned novel audiovisual associations either actively or passively. Results demonstrated that active learning enhanced subsequent unisensory item recognition and audiovisual associative recognition. Specifically, RTs during audiovisual associative and unisensory recognition were faster after active learning, as was accuracy during audiovisual associative recognition. Furthermore, motor-related regions (primary motor, primary somatosensory, and cerebellum) showed greater activation during the perception of actively learned audiovisual associations compared with passive learning. Additionally, brain regions implicated in audiovisual integration (e.g., STS) showed greater multisensory gain after active learning than after passive learning. Finally, functional connectivity between visual and motor cortices was stronger after active learning than passive learning (Butler et al., 2011).

The current study aims to expand upon this previous work to explore further the impact of active learning on subsequent perception and memory. An unexpected result from our previous work (Butler et al., 2011) was that, subsequent to active learning, unisensory visual and auditory items did not reactivate motor systems in the same ways that audiovisual presentation did. This finding is potentially in conflict with several other studies showing that unisensory information can lead to motor reactivation after active experience (De Lucia, Camen, Clarke, & Murray, 2009; James & Atwood, 2009; James & Mauoene, 2009; Etzel, Gazzola, & Keysers, 2008; Mutschler et al., 2007; Weisberg, van Turenhout, & Martin, 2007; James & Gauthier, 2006; Masumoto et al., 2006; Longcamp, Anton, Roth, & Velay, 2005; Grezes & Decety, 2002; Nyberg et al., 2001; Pulvermuller, Harle, & Hummel, 2001; Chao & Martin, 2000; Nilsson et al., 2000). Furthermore, in the previous study, there was enhanced visuo-motor connectivity during audiovisual presentation for actively learned associations, but not for unisensory items.

This lack of increased motor activation and enhanced visuo-motor connectivity during unisensory presentation for actively learned items might have occurred because of the nature of the actions in this previous study. Crucially, the actions were not specific to each item being studied. That is, in Butler et al. (2011), each object–sound pairing was associated with the same basic reaching/grasping/pressing action. This may have caused the individual objects and sounds to be less dissociated. This is important to consider because in real-world experiences we learn specific actions for specific objects even if the end product is the same. For example, typing and printing both result in

letter formation but involve very different actions based on the tool that we use to produce the output. To complicate things, even the same object may afford a different action depending on context: Holding a pitcher to pour liquid from it requires a different pattern of actions than holding a pitcher to hand it to someone else. Thus, learning the association between a percept and an action is very specific to the object and the situation. This specific coupling may lead to enhanced individuation of the object by the associated action, but if the action is not specific to the object, then the coupling may not occur—the action is general to many objects. The lack of individuation between unisensory items may explain why motor reactivation and enhanced visuo-motor connectivity did not occur during subsequent perception of actively learned unisensory items. Alternatively, this lack of motor reactivation and visuo-motor connectivity for unisensory items may suggest that motor reactivation is not incidental but rather is task specific. That is, the study session is multisensory and incorporates action but maybe motor systems are only reactivated during multisensory tasks, in this case, associative recognition.

To address this issue, the current study required the performance of a unique action during the active training condition that was associated with each object–sound relationship, in the hopes that this more ecologically valid way of acting on objects may serve to enhance the individuation of objects and sounds. Because each action object–sound pairing was unique, it was important to test perception and recognition of events, rather than of static images of objects. Therefore, we changed the test stimuli from static, 2-D images (as in Butler et al., 2011) to videos that depict the learned actions, rendering the visual recognition of action events rather than of static objects. Another important consequence of this is that the dynamic auditory stimuli could be better equated with the visual stimuli as compared with the previous study.

By implementing these changes, we are able to test the hypothesis that actions unique to object structure facilitate reactivation of motor systems during unisensory perception by contributing to the process of object individuation. In addition, we expanded upon our previous work by tracking learning during the training session to compare recognition immediately after learning and with a delay. This change allows us to discover whether or not active and passive learning are equally as efficient in terms of recognition. We maintained, however, our multisensory associative recognition tasks to replicate and extend our previous findings.

METHODS

Participants

Fifteen individuals (nine women and six men) participated in the study (mean age = 23 years, *SD* = 3 years). All participants gave informed consent according to the

guidelines of the Indiana University Institutional Review Board. All participants reported right handedness and normal or corrected-to-normal vision and no known neurological deficits. Participants were compensated for their time.

Stimuli

The stimuli presented during learning were 20 novel 3-D objects that each created a unique sound when a unique action was performed on or with them (see Figure 1 for stimuli examples and Supplementary Figure 1 for a list of all stimuli and associated actions). These novel 3-D objects were made of gray-colored lightweight ABS plastic using fused deposition modeling with a Stratasys Prodigy Plus (Eden Prairie, MN) rapid prototyping machine. Various sound producing parts were added to these objects so that they would produce unique sounds upon manual interaction. Each object had one specific associated unimanual or bimanual action that when performed created a sound (these were counterbalanced across active and passive conditions). For testing purposes, 2-sec duration black and white movies of the objects being acted upon from an egocentric perspective were created—termed events. Both congruent and incongruent audiovisual events were created. The congruent audiovisual stimuli matched the audio and video content from the actual objects used during training. The incongruent audio stimuli were movies with mismatched audio coupled with object movements. The incongruent stimuli were created such that the video of the object being acted upon was a plausible match to the mismatched sound. For example, one object required flicking a roller on its top that produced a “rattle” sound during learning. During test, the incongruent event displayed the same action on the same object producing a “clicking” sound instead of a “rattle.” In addition to these events, both unisensory visual only (the video of the movies without audio content) and unisensory audio (the audio of the movies without video content) were created. For the purpose of functional localization scrambled versions of the video only and audio only stimuli were created as well.

General Procedure

The experiment consisted of two sections, a learning/testing session and an fMRI scan session. First, participants underwent a learning/testing session and an fMRI session. The learning portion of the learning/testing session involved training participants to move an object, which then produced a unique sound or having participants watch an experimenter move the objects. The test portion of the learning/testing session was interleaved between each round of learning. Testing involved recognizing videos of an experimenter performing the learned actions on the learned objects. The videos were either congruent—the correct learned audiovisual association—or incongruent, the learned action producing a different sound. During these tests, participants respond yes or no whether the association was one that they learned. The learning/testing session lasted 45 min. Immediately after the learning/test session, participants underwent an fMRI scan session that included several types of runs (see procedural details below and Figure 2 for an illustration of these runs). The fMRI scan session lasted approximately 1 hr.

Learning/Testing Session

Each participant learned half of the objects/actions actively and half passively. During active learning, the participants acted on the objects to make associated sounds, and during passive learning, participants viewed an experimenter act on the objects. The visual and auditory experience of the participants during the active and passive conditions was kept as similar as possible. The only difference was that during the passive condition the participants watched the experimenter act on/with the objects. To keep the visual perception of the two groups as similar as possible, care was taken to ensure that the perspective was equated by having the experimenter and participant sit next to each other. The objects were acted upon in the space front of and between the participant and experimenter. This allowed for the participants to see the objects being acted upon from an egocentric point of view in both the active and passive conditions. Furthermore, the experimenter

Figure 1. Stimuli examples. Photos of the novel sound-producing 3-D objects used during active and passive training.

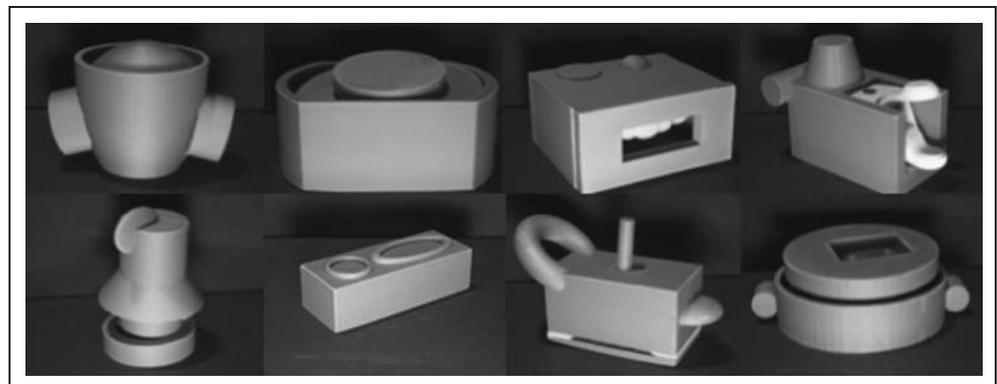
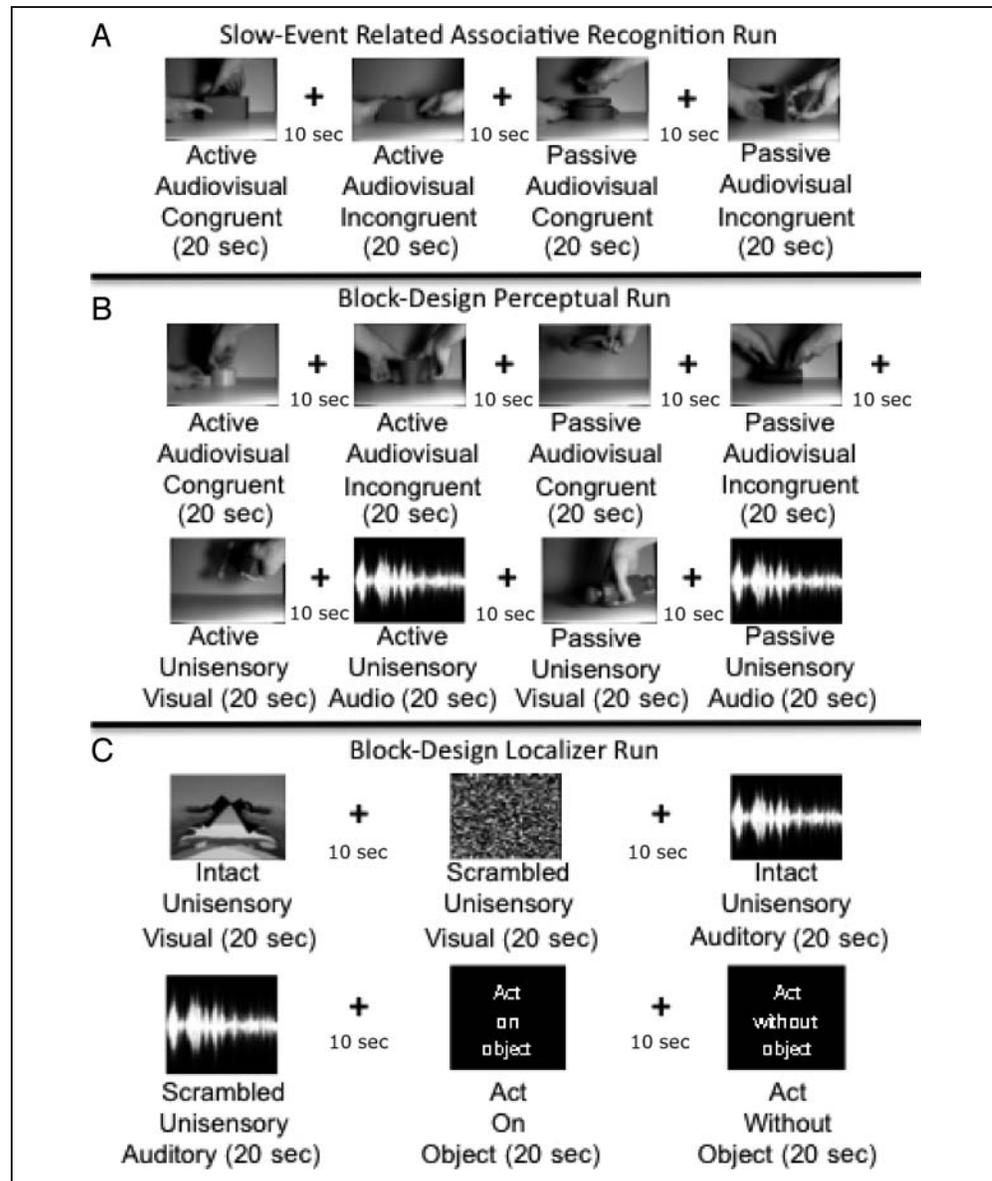


Figure 2. fMRI run design. Illustration of the design of the three different types of fMRI runs.



placed the object before the participant in the same orientation before interaction for both the active and passive conditions. This orientation matched what participants were presented with subsequently in the videos they saw in the MRI environment. The action events in the videos all started at these same orientations. All 20 objects were presented exactly at five different times, and the order of presentation was randomized. Each time they were presented with an object, they acted on it or watched it acted upon for three times. Therefore, all participants acted on or watched each object acted upon for a total of 15 times over the course of the training session.

Because the current design is more like everyday interaction with objects than previous work, we were concerned that it may be very difficult to learn these associations simply through observation (passive learning). To this end, we were careful to measure the amount of time to

learn associations in both conditions. If active and passive learning reached the same criterion, we can say with more confidence that differences were due to the different types of learning experiences as opposed to differences in the degree of learning. Therefore, associative recognition tests were given after each of the rounds of active/passive learning. Overall, five associative recognition tests were given across the whole learning session for each participant. During each associative recognition test, the participants were presented with audiovisual movies showing each object being acted upon to produce its associated sound. There were 40 movies in total; each was 2 sec in duration followed by accuracy feedback (correct or incorrect) for 1 sec with 1 sec more of fixation before the next trial. Half of the movies depicted actively learned objects and half passively learned objects. Half of the active movies were congruent, and half were incongruent. Similarly, half

of the passive movies were congruent, and half were incongruent. Congruent movies matched the audiovisual associations learned, whereas incongruent movies were mismatches of the video and audio of the same stimuli. Participants responded with the index or middle finger of their right hand, deciding whether the movies were congruent or incongruent. The learning/testing session lasted 45 min.

fMRI Procedure

Immediately after the learning/testing session, participants were brought to the imaging research facility. After instructions and safety screening, participants underwent the imaging session that lasted between 1 hr and 1 hr 30 min. Functional imaging was divided into eleven runs of three different types (see Figure 2). There was one “associative recognition” run, eight “perceptual runs,” and two functional localizer runs. After the functional runs were complete, an anatomical series was collected.

Associative Recognition Run

During the slow event-related associative recognition, run participants performed an associative recognition test identical to those given during the learning session (see Figure 2A). This testing was important to check whether participants could still correctly recall audiovisual associations above chance in the MRI environment. Also it was important to test if any differences in associative recognition accuracy were present between actively and passively learned associations after the delay between training and the scanning session. A total of four types of audiovisual movies were presented: active audiovisual congruent, active audiovisual incongruent, passive audiovisual congruent, and passive audiovisual incongruent. The task was for participants to use the index and middle finger of their right hand to indicate whether the stimuli were congruent or incongruent. Each movie was 2 sec in duration followed by 10 sec of fixation. There were 10 movies in each condition for a total of 40 movies. The total length of this run type was 8 min and 20 sec. As stated, this run served to inform whether or not performance on associative recognition outside of the scanning environment was the same within the scanning environment after the delay; however, it also served to provide an index of neural activation patterns during such a task.

Perceptual Runs

During the eight blocked perceptual imaging runs, participants viewed blocks that were 20 sec each, containing 10 videos of 2-sec duration that depicted one of eight conditions: active visual only, active audio only, active audiovisual congruent, active audiovisual incongruent, passive visual only, passive audio only, passive audiovisual congru-

ent, and passive audiovisual incongruent (see Figure 2B). All conditions were presented within each run in different random orders, and this was counterbalanced. Each condition was therefore seen eight times. The instruction before the runs was to pay attention to the stimuli that would be presented, and participants passively viewed the stimuli. The participants did not make visible hand movements as confirmed by a camera in the bore of the magnet. The total length of each of these runs was 4 min and 20 sec.

Functional Localizer Runs

We also included extensive individual functional localizers to account for individual differences in the spatial location of crucial perceptual and motor regions (see Figure 2C). These changes were important for us to understand better how object-specific actions affect multisensory integration and how this experience affects processing of events—a more ecological situation than processing of static images. Overall, the current study allows us to probe how active learning of unique novel audiovisual associations impacts subsequent perception and recognition at the behavioral and neural levels.

The blocked design localizer runs were performed to acquire several brain ROIs for in-depth analyses. Conditions included unisensory visual events and unisensory auditory sound presentation of both actively and passively learned stimuli as well as scrambled versions of both these movies and sounds. Additionally, there were two different types of blocks that were used to acquire recruitment of motor systems. During one block, participants were required to physically manipulate an object in the scanner using both hands, and in the other, they had to perform the same movements with both hands but without the object. The total length of each of this run type was 5 min 20 sec.

Functional Imaging Parameters

Imaging was performed using a 3-T Siemens Magnetom Trio (Munich, Germany) whole-body MRI system (with a TIM upgrade) and a 32-channel radio-frequency 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 Superlab software via an Apple Macbook laptop.

The field of view was $22 \times 22 \times 9.9$ 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 $1.7 \text{ mm} \times 1.7 \text{ mm} \times 3.4 \text{ mm}$. Images were acquired using an echo-planar technique (echo time = 28 msec, repetition time = 2000 msec, flip angle = 70°) for BOLD-based imaging. High-resolution T1-weighted

anatomical volumes were acquired using a 3-D Turbo-flash acquisition.

fMRI Data Analysis

BrainVoyager QX 2.2.0 (Brain Innovation, Maastricht, Netherlands) was used to analyze the fMRI data. Preprocessing included slice scan-time correction, 3-D motion correction, Gaussian spatial smoothing (6 mm), and linear trend removal. Individual functional volumes were coregistered to anatomical volumes with an intensity-matching, rigid body transformation algorithm. Individual anatomical volumes were normalized to the stereotactic space of Talairach and Tournoux (1988) using an eight-parameter affine transformation, with parameters selected by visual inspection of anatomical landmarks. Applying the same affine transformation to the coregistered 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.

Functional ROI analyses were performed using regions defined at the level of the individual participant (Saxe, Brett, & Kanwisher, 2006). For these whole-brain analyses in individuals, the data from the localizer runs were entered into general linear models using an assumed two-gamma hemodynamic response function. The data from both localizer runs was concatenated in each subject, and the beta values calculated from all conditions per subject. The baseline was the average of the rest intervals across both runs. The independent functional localizer was used to create ROIs in motor, somatosensory, cerebellum, visual (lateral occipital complex [LOC]), auditory (Heschl's gyrus), and audiovisual (STS) regions. Motor regions were found using the contrast: acting on object + acting without object > fixation (weighted and balanced). Somatosensory and cerebellar regions were found using the contrast: acting on object > acting without objects. This contrast was used to localize somatosensory regions because when acting on the object participants were receiving haptic input but when acting without objects this haptic input was absent. Presumably due to this difference, the contrast reliably activated bilateral somatosensory regions (in the post central gyrus). This contrast also reliably activated bilateral cerebellar regions in individuals and was thus used to identify these regions as well. The visual (LOC) ROIs were created using the contrast: visual intact > visual scrambled. This is a common way of isolating the object-specific LOC (James, Culham, Humphrey, Milner, & Goodale, 2003). The auditory (Heschl's gyrus) ROIs were created using the contrast: auditory intact > fixation. We attempted to define auditory regions by contrasting intact and scrambled audio, but the resultant individual ROIs were not reliably found across individuals. Finally, the audiovisual (STS) ROIs were made using the conjunction of visual intact

and auditory intact conditions. Previous work has shown that the STS, when defined with this conjunction, shows properties of multisensory integration such as inverse effectiveness, spatial congruency, and temporal synchrony (Stevenson, Altieri, Kim, Pisoni, & James, 2010; Stevenson & James, 2009; Stevenson, Geoghegan, & James, 2007). These individual ROIs are shown in Figure 3. Tables 1 and 2 show specific cluster- and statistical-related information about all individual ROIs.

Functional connectivity was assessed using the RFX Granger Causality Mapping v2.5 plugin in BrainVoyager. Seed regions for the analysis were created using each individual's left and right visual ROIs (LOC). Instantaneous correlations were calculated for BOLD activation produced during the main runs. 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 within-groups ANCOVA 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.

RESULTS

Behavioral Results

Before fMRI scanning, associative recognition was measured across five time points. Active and passive conditions were compared at each time point using pairwise *t* tests.

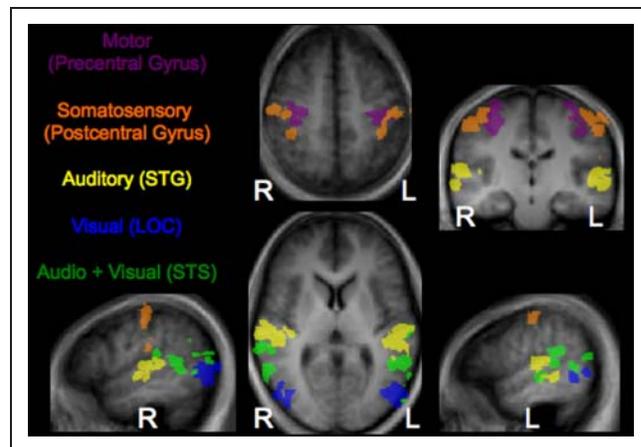


Figure 3. Individual functional ROIs. Whole-brain map showing individual functional ROIs derived from independent localizer runs. Specifically, each individual participants cluster for each of the different types of functional localizer is shown. The different types of functional localizers are shown in different colors (each of the individuals have the same color for each type). Colors were coded in the following way: motor (precentral gyrus) in purple, somatosensory (postcentral gyrus) in orange, auditory (superior temporal gyrus) in yellow, visual (lateral occipital gyrus) in blue, and the conjunction of audio and visual (STS) in green.

Table 1. Perceptual-related Individual ROI Cluster Information

Participant No.	Left				Right			
	Cluster Size	Talairach Coordinates for Peak (x, y, z)	Peak t Value	Peak p Value	Cluster Size	Talairach Coordinates for Peak (x, y, z)	Peak t Value	Peak p Value
<i>A. Visual (LOC)</i>								
1	605	(-55, -68, 3)	9.54	<.000001	635	(47, -62, 0)	7.58	<.000001
2	506	(-49, -80, 9)	13.64	<.000001	520	(44, -68, 6)	10.53	<.000001
3	512	(-46, -68, 3)	11.63	<.000001	533	(41, -74, 9)	7.99	<.000001
4	574	(-43, -71, 6)	12.93	<.000001	676	(41, -68, 3)	14.12	<.000001
5	584	(-46, -74, 3)	12.79	<.000001	480	(44, -83, -3)	9.4	<.000001
6	437	(-49, -71, -3)	11.47	<.000001	483	(41, -71, -12)	7.66	<.000001
7	430	(-49, -68, 15)	12.09	<.000001	608	(50, -50, 0)	9.89	<.000001
8	560	(-46, -65, 6)	12.41	<.000001	543	(38, -74, 3)	14.36	<.000001
9	598	(-49, -60, 3)	13.63	<.000001	513	(41, -71, -3)	11.31	<.000001
10	478	(-55, -71, 12)	12.42	<.000001	442	(44, -76, 13)	12.08	<.000001
11	632	(-43, -74, 3)	14.07	<.000001	580	(41, -74, 15)	10.34	<.000001
12	531	(-41, -70, 9)	19.82	<.000001	576	(41, -74, 0)	14.09	<.000001
13	476	(-49, -68, 6)	12.14	<.000001	515	(59, -53, 6)	5.9	<.000001
14	592	(-49, -65, 6)	12.04	<.000001	449	(47, -65, 6)	13.85	<.000001
15	678	(-52, -71, 3)	10.63	<.000001	514	(44, -74, -3)	14.54	<.000001
<i>B. Auditory (STG)</i>								
1	560	(-55, -20, 6)	7.52	<.000001	749	(62, -26, 8)	7.42	<.000001
2	575	(-46, -20, 3)	11.01	<.000001	433	(35, -23, 12)	8.22	<.000001
3	475	(-46, -23, 9)	15.46	<.000001	636	(56, -16, 12)	12.96	<.000001
4	536	(-43, -29, 12)	14.83	<.000001	633	(59, -20, 9)	13.52	<.000001
5	549	(-40, -29, 9)	14.45	<.000001	479	(62, -32, 15)	15.41	<.000001
6	460	(-52, -17, 9)	14.01	<.000001	528	(62, -26, 0)	11.72	<.000001
7	467	(-52, -32, 9)	11.76	<.000001	634	(56, -20, 12)	11.25	<.000001
8	455	(-52, -14, 0)	15.42	<.000001	503	(59, -14, 6)	14.48	<.000001
9	517	(-52, -23, 9)	15.43	<.000001	578	(56, -23, 9)	19.42	<.000001
10	507	(-37, -32, 15)	7.37	<.000001	537	(59, -26, 15)	16.34	<.000001
11	518	(-40, -32, 15)	15.94	<.000001	502	(59, -20, 9)	13.88	<.000001
12	496	(-46, -41, 12)	15.81	<.000001	535	(62, -26, 0)	16.84	<.000001
13	631	(-52, -29, 9)	3.59	<.000001	590	(47, -23, 6)	8.1	<.000001
14	432	(-52, -23, 9)	17.89	<.000001	399	(53, -20, 9)	17	<.000001
15	741	(-55, -20, 0)	6.52	<.000001	549	(62, -17, 0)	8.86	<.000001
<i>C. Audiovisual (STS)</i>								
1	638	(-55, -68, 15)	3.83	<.00015	436	(56, -62, 12)	3.24	<.00131
2	503	(-61, -38, 12)	3.17	<.00166	487	(53, -38, 21)	2.26	<.02428

Table 1. (continued)

Participant No.	Left				Right			
	Cluster Size	Talairach Coordinates for Peak (x, y, z)	Peak t Value	Peak p Value	Cluster Size	Talairach Coordinates for Peak (x, y, z)	Peak t Value	Peak p Value
3	525	(-46, -29, 15)	13.78	<.000001	539	(56, -30, 15)	9.87	<.000001
4	649	(-52, -47, 12)	7.19	<.000001	445	(59, -35, 15)	9.3	<.000001
5	689	(-52, -23, 12)	5.45	<.000001	722	(53, -26, 0)	3.93	<.0001
6	526	(-55, -44, 15)	4.66	<.000004	555	(69, -23, 15)	6.33	<.000001
7	500	(-49, -44, 9)	6.76	<.000001	457	(56, -32, 12)	7.75	<.000001
8	525	(-58, -62, 33)	5.25	<.000001	621	(56, -35, 18)	6.91	<.000001
9	680	(-55, -50, 6)	4.79	<.000002	528	(56, -32, 21)	6.54	<.000001
10	-	-	-	-	688	(62, -41, 21)	3.47	<.000589
11	662	(-55, -74, 18)	5.94	<.000001	619	(41, -62, 27)	6.11	<.000001
12	474	(-55, -62, 3)	11.35	<.000001	666	(56, -53, 9)	9.65	<.000001
13	523	(-56, -77, 12)	2.59	<.001	-	-	-	-
14	664	(-46, -41, 17)	5.62	<.000001	558	(44, -32, 27)	10.5	<.000001
15	668	(-58, -41, 6)	4.67	<.000004	469	(62, -35, 9)	3.92	<.000107

Relevant cluster-related information concerning individual functional ROIs in visual (LOC), auditory (STG), and audiovisual (STS) perceptual-related regions.

Table 2. Motor-related Individual ROIs Cluster Information

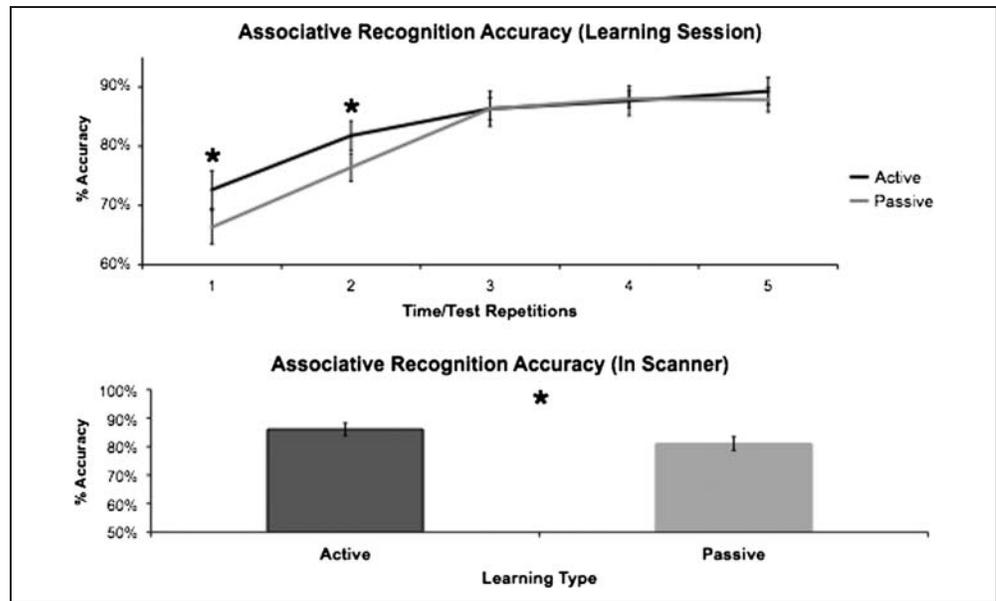
Participant No.	Left				Right			
	Cluster Size	Talairach Coordinates for Peak (x, y, z)	Peak t Value	Peak p Value	Cluster Size	Talairach Coordinates for Peak (x, y, z)	Peak t Value	Peak p Value
<i>A. Motor (precentral gyrus)</i>								
1	559	(-31, -23, 48)	13.32	<.000001	585	(29, -17, 51)	13.72	<.000001
2	511	(-37, -26, 51)	15.72	<.000001	510	(35, -22, 51)	11.39	<.000001
3	507	(-28, -23, 54)	20.08	<.000001	673	(29, -25, 57)	20.12	<.000001
4	481	(-37, -20, 48)	20.81	<.000001	516	(23, -14, 57)	8.06	<.000001
5	470	(-31, -14, 63)	20.83	<.000001	573	(32, -23, 63)	24.92	<.000001
6	483	(-46, -26, 48)	12.86	<.000001	577	(32, -29, 45)	13.42	<.000001
7	460	(-34, -23, 48)	4.18	<.000046	579	(29, -20, 51)	12.59	<.000001
8	450	(47, -20, 51)	24.31	<.000001	449	(41, -23, 57)	24.69	<.000001
9	468	(-31, -20, 72)	23.38	<.000001	454	(33, -30, 57)	15.97	<.000001
10	457	(-37, -11, 57)	11.74	<.000001	507	(32, -20, 48)	12.3	<.000001
11	466	(-40, -14, 54)	23.54	<.000001	517	(35, -20, 45)	31.04	<.000001
12	546	(-40, -17, 42)	20.53	<.000001	589	(32, -14, 42)	16.74	<.000001
13	528	(-34, -32, 52)	13.07	<.000001	616	(32, -29, 54)	11.77	<.000001
14	634	(-34, -26, 48)	18.06	<.000001	620	(32, -20, 66)	21.06	<.000001
15	546	(-37, -14, 57)	28.93	<.000001	604	(35, -17, 60)	26.05	<.000001

Table 2. (continued)

Participant No.	Left				Right			
	Cluster Size	Talairach Coordinates for Peak (x, y, z)	Peak t Value	Peak p Value	Cluster Size	Talairach Coordinates for Peak (x, y, z)	Peak t Value	Peak p Value
<i>B. Haptic (postcentral gyrus)</i>								
1	550	(-52, -17, 45)	5.84	<.000001	547	(35, -23, 63)	8.11	<.000001
2	606	(-34, -35, 51)	6.53	<.000001	459	(35, -38, 45)	4.76	<.000003
3	565	(-55, -14, 51)	12.01	<.000001	469	(53, -17, 61)	12.19	<.000001
4	607	(-46, -23, 57)	7.85	<.000001	639	(38, -26, 45)	7.28	<.000001
5	574	(-46, -20, 51)	14.74	<.000001	490	(38, -14, 62)	15.88	<.000001
6	554	(-53, -23, 37)	2.96	<.00033	576	(50, -20, 45)	2.51	<.0127
7	408	(-36, -42, 54)	3.94	<.000116	547	(32, -41, 54)	8.88	<.000001
8	456	(-55, -20, 48)	13	<.000001	570	(53, -17, 48)	13.19	<.000001
9	465	(-40, -23, 69)	13	<.000001	574	(44, -32, 63)	18.89	<.000001
10	668	(-58, -23, 58)	8.46	<.000001	448	(47, -17, 48)	8.35	<.000001
11	433	(-49, -20, 48)	10.09	<.000001	676	(43, -20, 51)	8.93	<.000001
12	476	(-49, -23, 24)	8.73	<.000001	546	(32, -26, 39)	5.58	<.000001
13	638	(-40, -26, 45)	7.86	<.000001	533	(35, -38, 51)	5.84	<.000001
14	585	(-46, -17, 54)	13.4	<.000001	590	(50, -20, 51)	15.72	<.000001
15	492	(-43, -14, 57)	12.31	<.000001	664	(41, -17, 54)	12.51	<.000001
<i>C. Cerebellum</i>								
1	544	(-16, -47, -12)	16.76	<.000001	565	(14, -44, -12)	15.56	<.000001
2	582	(-22, -44, -21)	14.21	<.000001	508	(23, -50, -18)	12.82	<.000001
3	645	(-13, -53, -12)	16.74	<.000001	605	(11, -53, -12)	15.77	<.000001
4	448	(-25, -56, -18)	8.79	<.000001	447	(14, -56, -15)	11.42	<.000001
5	522	(-25, -47, -24)	18.89	<.000001	451	(23, -47, -24)	18.9	<.000001
6	466	(-16, -50, -15)	16.6	<.000001	535	(11, -50, -21)	15.1	<.000001
7	668	(-19, -53, -12)	18.46	<.000001	589	(14, -50, -12)	15.23	<.000001
8	542	(-19, -47, -15)	23.7	<.000001	569	(17, -44, -18)	22.82	<.000001
9	600	(-28, -41, -21)	21.14	<.000001	501	(26, -53, -18)	13.69	<.000001
10	461	(-28, -50, -18)	12.78	<.000001	501	(23, -50, -21)	13.13	<.000001
11	578	(-19, -41, -18)	25.5	<.000001	601	(11, -41, -15)	25.23	<.000001
12	537	(-16, -50, -15)	21.58	<.000001	566	(8, -50, -9)	17	<.000001
13	439	(-22, -41, -18)	11.78	<.000001	463	(20, -35, -18)	9.86	<.000001
14	458	(-19, -44, -21)	18.02	<.000001	554	(17, -41, -21)	14.98	<.000001
15	549	(-13, -50, -12)	21.19	<.000001	447	(14, -53, -15)	23.23	<.000001

Relevant cluster-related information concerning individual functional ROIs in motor (precentral gyrus), somatosensory (postcentral gyrus), and cerebellar motor-related regions.

Figure 4. Behavioral results. Behavioral accuracy of audiovisual associative recognition during both the learning (top) and scanning sessions (bottom). Active learning showed increased accuracy during the initial two time points during the learning session. During the scanning session, actively learned associations had greater associative recognition accuracy. *Statistically significant difference at $p < .05$ for all graphs. Error bars represent standard error of the mean.



Active learning demonstrated greater associative recognition accuracy than passive learning at both the first, $t(14) = 1.85, p = .043$, and second, $t(14) = 1.96, p = .035$, time points (see Figure 4). Importantly both active and passive learning reached the same degree of accuracy at the

third, fourth, and fifth time points. These results demonstrate that active interaction speeded learning during the session, but both active and passive learning reached the same criterion by the end of training. After this training, associative recognition was measured in the scanner once

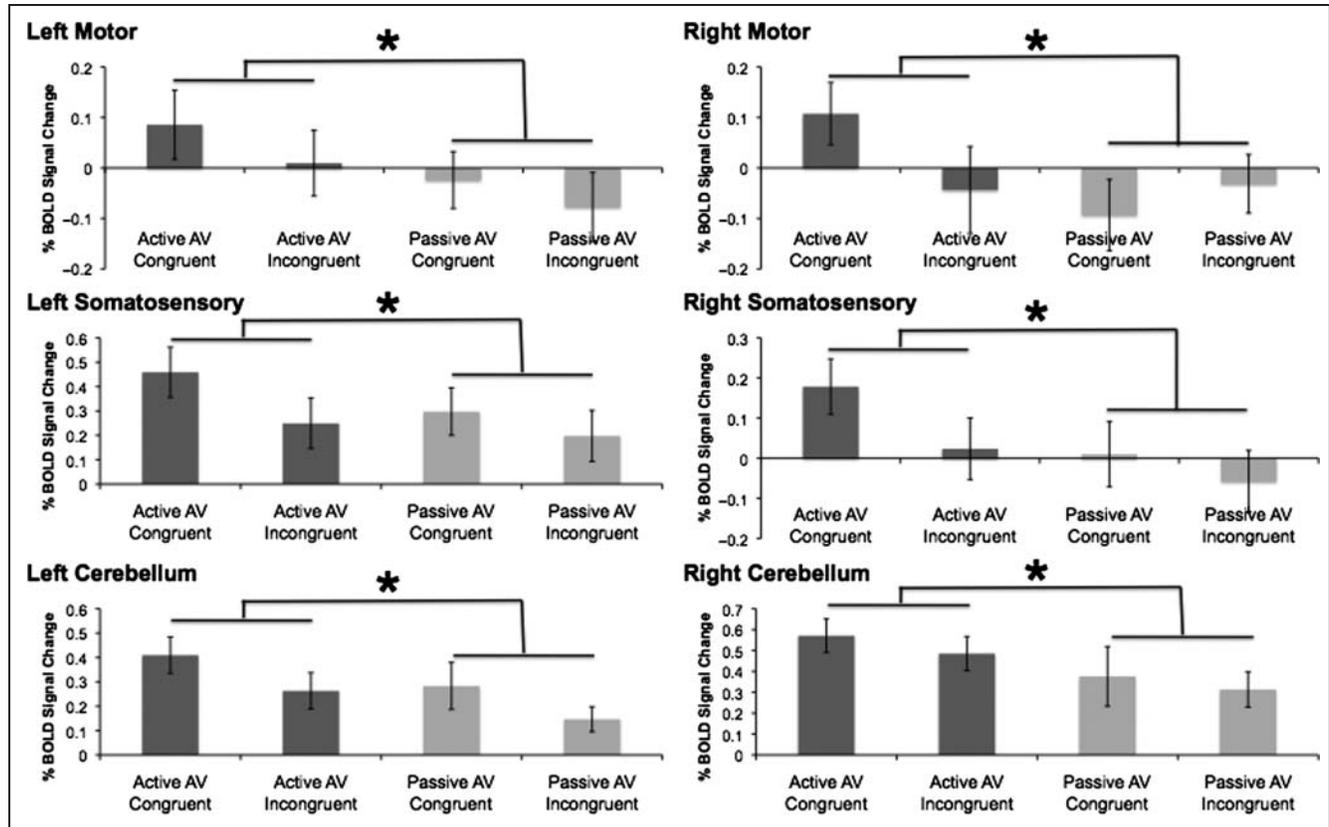


Figure 5. Functional ROI results for perceptual runs. Graphs show all functional ROIs with a significant main effect of learning type in which active learning showed greater activation overall compared with passive learning during perceptual runs. *Statistically significant difference at $p < .05$ for all graphs. Error bars represent standard error of the mean.

more. The behavioral data from the associative recognition test given in the scanner showed that active learning had significantly greater accuracy than passive learning, $t(14) = 1.94, p = .036$ —a result that differed from our last behavioral measure outside of the scanner. The significance of this different result is discussed below.

Perceptual Runs Functional ROI Results

Twelve 2 (active, passive) \times 2 (congruent, incongruent) repeated-measures ANOVAs were performed on the data from the perceptual runs that was gathered from the functionally defined ROIs. Results from this analysis demonstrated that active learning produced significantly greater activation than passive learning in sensori-motor systems, including the bilateral motor ROIs [left motor: $F(1, 14) = 4.4, p = .05$; right motor: $F(1, 14) = 5.14, p = .04$], bilateral somatosensory ROIs [left somatosensory: $F(1, 14) = 8.4, p < .01$; right somatosensory: $F(1, 14) = 7.43, p = .02$], and bilateral cerebellar ROIs [left cerebellar: $F(1, 14) = 9.29, p = .01$; right cerebellar: $F(1, 14) = 11.91, p = .004$; see Figure 5]. It should be noted that these effects were not found during unisensory auditory or unisensory visual presentation in the perceptual runs. Furthermore, there were no such effects for the perceptual run data in visual, auditory, or audiovisual ROIs.

Overall for the perceptual runs, activation was greater in sensori-motor-related regions to actively compare with passively learned associations during audiovisual presentations, but no differences were seen between active and passively learned associations in perceptual regions. Greater activation for actively than passively learned associations in motor-related regions replicates our previous study (Butler et al., 2011). However, in our previous study, an interaction between learning type and congruency was present, showing greater activation overall for congruent actively learned associations. In the current results, there was a main effect of learning type in motor-related regions showing that overall actively learned associations showed stronger activation than passively learned associations. Additionally, greater activation for motor-related regions was bilateral in the current study but tended to be unilateral (left lateralized for motor and haptic regions and right lateralized for cerebellar regions) in our previous work.

Recognition Runs Functional ROI Results

Twelve 2 (active, passive) \times 2 (congruent, incongruent) repeated-measures ANOVAs were performed on the data from the recognition runs in the functionally defined ROIs. Incorrect trials were excluded from analysis. Active learning showed significantly greater activation than passive learning in a subset of sensori-motor-related regions that also showed this effect during the perceptual runs (see above). Specifically, this effect was seen in the left motor ROI, $F(1, 14) = 7.31, p = .017$, the left somatosensory ROI, $F(1, 14) = 7.96, p = .014$, and the right cerebellar ROI,

$F(1, 14) = 5.33, p = .037$ (see Figure 6). These effects were not found during unisensory auditory or unisensory visual presentation during the recognition run. Furthermore, there were no such effects for the recognition run data in visual, auditory, or audiovisual ROIs.

Overall, similar to the perceptual runs, activation was greater in sensori-motor-related regions to actively compare with passively learned associations during audiovisual associative recognition, but no differences were seen between active and passively learned associations in perceptual regions. This finding is novel, as our previous experiment (Butler et al., 2011) did not test for differences between active and passive learning during a recognition task.

Whole-brain Functional Connectivity

The results of a whole-brain functional connectivity analyses using seed regions in bilateral visual object-selective LOC revealed a right lateralized premotor region showing a stronger instantaneous correlation for the active than passive learning condition during audiovisual perception (see Figure 7). This right lateralized premotor region was highly overlapping for both the left and right LOC seed regions. For the left LOC, the Talairach coordinates for

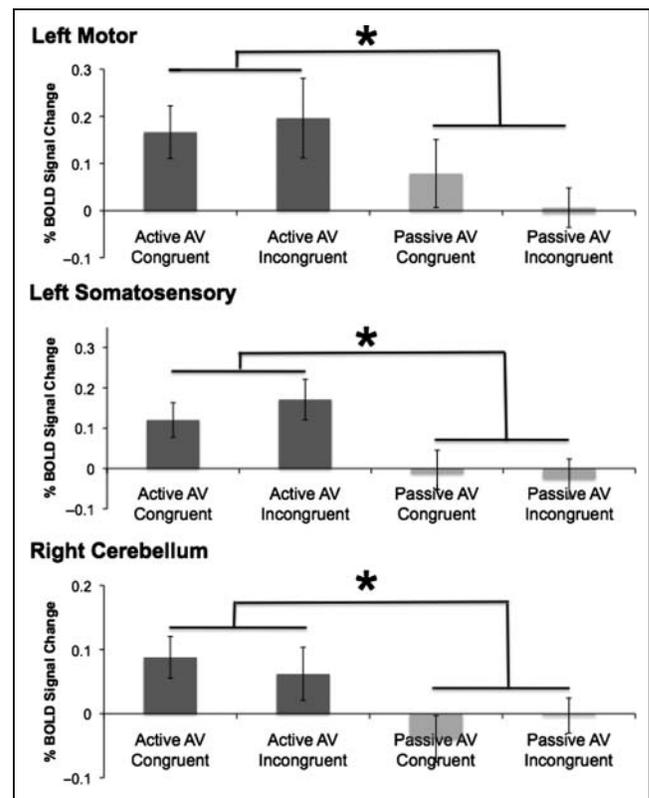
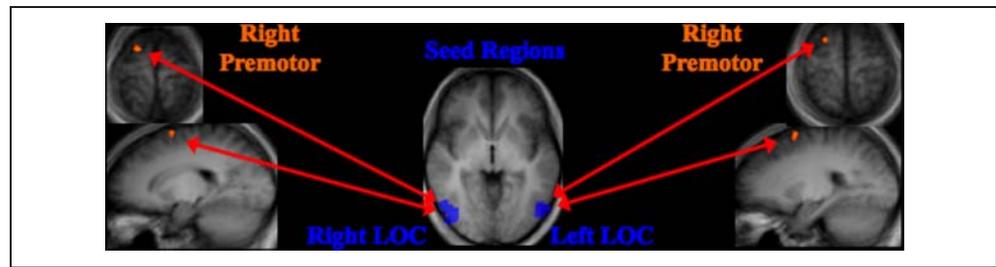


Figure 6. Functional ROI results for recognition run. Graphs show all functional ROIs in which a significant main effect of learning type in which active learning showed greater activation overall compared with passive learning during the recognition run. *Statistically significant difference at $p < .05$ for all graphs. Error bars represent standard error of the mean.

Figure 7. Whole-brain functional connectivity results. The seed regions, shown in blue, are in the left and right LOC (based on individual ROIs). Whole-brain correlational analysis for both of these seed regions showed a stronger correlation after active than passive learning with a highly overlapping right premotor region (shown in orange).



the peak of this correlated right premotor region was $x = 20, y = 7, z = 59$. For the right LOC, the Talairach coordinates for the peak of this correlated right premotor region was $x = 18, y = 5, z = 62$. While, as in Butler et al. (2011), motor-related and visual processing regions showed a greater correlation after active learning, the precise motor and visual regions correlated were different. Butler et al. (2011) found that a left-lateralized primary motor region showed a stronger correlation with several visual processing regions after active learning. However, in the current study a more posterior right-lateralized premotor region was correlated with visual processing regions.

DISCUSSION

The results of the current study demonstrated that active learning impacted subsequent perception and recognition in five significant ways. First, during training, active experience led to faster audiovisual associative learning than passive experience. Second, actively learned associations were also better retained after a delay. Third, active learning was associated with greater neural recruitment in motor and haptic processing-related regions (precentral gyrus, postcentral gyrus, cerebellum) during both subsequent perception and recognition tasks. Fourth, differences in learning conditions can affect perception of events, not only perception of static images. Finally, functional connectivity between object-selective visual regions (LOC) and premotor regions was enhanced during the presentation of actively learned audiovisual associations when compared with passively learned associations. Crucially, for the motivation of the current work, motor-related reactivation effects and enhanced visuo-motor connectivity only occurred when participants were subsequently presented with actively learned audiovisual associations. Such effects did not occur during the unisensory presentation of actively learned visual and auditory items. As discussed below, this has important implications for the nature of these effects and understanding how the retrieval of motor information is related to subsequent perception and memory.

Active Learning Results in Faster Audiovisual Associative Recognition than Passive Learning

In a general sense, the current results of the enhancement in behavioral accuracy for audiovisual learning and recog-

nition performance replicates and extends previous work focusing on the effects of active learning on unisensory information. Past studies show that actively exploring novel unisensory visual objects leads to faster RTs during later visual item recognition (Sasaoka et al., 2010; James et al., 2002; Harman et al., 1999). Extending this previous work and the work of Butler et al. (2011), the current study provides the first evidence that active manipulation of sound-producing objects speeds audiovisual associative learning. Specifically, during training, active experience led to faster audiovisual associative learning than passive experience. Therefore, the current study provides novel findings relating to the behavioral impact of active learning.

Active Learning Facilitates Retention of Information

Another novel behavioral finding of the current work relates to the affect of active learning on the retention of information over delay. Our previous work, Butler et al. (2011), showed that active learning enhanced both unisensory item recognition and audiovisual associative recognition. The current study replicates and extends this work by showing that active learning enhanced associative recognition accuracy after a delay whereas passive learning did not. That is, active and passive associations were initially learned to the same degree in the current work, but after a delay, active learning led to increased associative recognition, suggesting that this particular type of learning results in better retention. Furthermore, actively learned information was better retained and/or accessed even in the different and more noisy environment of the MRI scanner.

Active Learning Results in Recruitment of Motor- and Haptic-related Processing Regions during Subsequent Perception and Recognition

Reactivation of motor-related regions has previously been shown to occur during perception and recognition of unisensory information after active learning. Motor-related regions reactivated in such studies include primary motor (James & Swain, 2011; James & Atwood, 2009; James & Mauoene, 2009; Masumoto et al., 2006; Grezes & Decety, 2002; Senkfor, Petten, & Kutas, 2002; Nyberg et al., 2001; Nilsson et al., 2000), premotor regions (De Lucia et al., 2009; James & Mauoene, 2009; Etzel et al., 2008; Weisberg et al.,

2007; Longcamp et al., 2005; Chao & Martin, 2000), supplementary motor area (Grezes & Decety, 2002), insula (Mutschler et al., 2007), and the cerebellum (Imamizu, Kuroda, Miyauchi, Yoshioka, & Kawato, 2003; Nyberg et al., 2001). Additionally, active learning involves proprioceptive and haptic information. Haptic exploration during learning enhances the recognition of audiovisual associations (Fredembach, de Boisferon, & Gentaz, 2009), and fMRI studies have shown that somatosensory regions reactivate during the retrieval of haptic-related information (Stock, Roder, Burke, Bien, & Rosler, 2009).

Butler et al. (2011) demonstrated that active learning of audiovisual associations lead to reactivation of motor and haptic regions during subsequent perception of static objects. This previous study was not, however, able to test for this effect during actual recognition. The current study therefore extends the previous work by being the first to demonstrate that motor- and haptic-related regions reactivate to a greater degree after active than passive learning during both perception and recognition. Importantly, in the present study, unisensory visual or auditory perception did not lead to these relative reactivations after active learning. Motor- and haptic-related reactivations only occurred when actively learned congruent associations or incongruent pairings of actively learned items were subsequently perceived or recognized. These reactivations may be related to the behavioral enhancements in associative recognition after active learning. Previous work has suggested that some visual objects potentiate actions in an automatic fashion based on the presence of perceptual affordances (physical characteristics that suggest possible actions), thus leading to motor reactivation (Grezes, Tucker, Armony, Ellis, & Passingham, 2003; Grezes & Decety, 2002; Gibson, 1977). The current results suggest that, although affordances may automatically potentiate actions, such effects may only occur, or are at least enhanced, after motor interaction during learning. Actions, therefore, may be associated to perceptual affordances through self-generated action as opposed to perception of action.

Several theories propose that entities and events are stored as embodied representations that include access to the pattern of perceptual and motor activity occurring during learning (e.g., Fuster, 2009; Barsalou, 1999; Damasio, 1989). These theories would predict the reactivation of motor and haptic systems after active learning. Whereas the current study generally supports these theories, it also suggests that such theories should consider how associations between multiple items or types of information, such as the audiovisual associations in the current study, may be needed for or modulate reactivation. In the current study, the active condition consisted of associating at least three types of information: visual, auditory, and motor/haptic. During the perception and recognition of the actively learned audiovisual associations in which both the visual and auditory information was provided, there was greater activation in motor-related regions. However, this was not the case when only visual or only auditory information

was presented. Therefore, certain conjunctions of stimuli may be needed for reactivation to occur or at least may relatively enhance reactivation.

The now common finding of reactivation of motor-related systems to unisensory items upon recognition might be due to the direct association of motor experience with unisensory items. Apparently at odds with such a finding, in the current study, we found that motor-related reactivation only occurred in response to associations as opposed to items. This suggests that the type of task—item recognition versus associative recognition—may be important for how the brain later recruits motor-related information. In the current study, participants may have used motor-related information only when associations were presented because it aided them in determining whether an association was congruent or not. Motor information would not be useful in this way when unisensory items were presented. Overall, the evidence from the current results supports the view that the retrieval of motor information may be task specific—in this case, only occurring during audiovisual associative perception and recognition.

Perception of Actions versus Static Images

Our perceptions are not static, but change as we move and as others move. We designed this study to mimic this type of active interaction in the world, and therefore, also tested recognition and neural activation to dynamic images. Therefore, it is important to consider how subsequent perception of actions may be modulated by active experience. A recent meta-analysis of 104 functional neuroimaging experiments suggested that the observation of actions was associated with bilateral activation of regions in frontal, parietal, and posterior temporal cortex (Caspers, Zilles, Laird, & Eickhoff, 2010). Additionally, the possible role of the “human mirror system” (HMS) is important to consider in light of the design and results of the current study. Previous work has demonstrated that certain cells in frontal, STS, and parietal regions of the nonhuman primate, termed mirror neurons, fire during both the performance and the observation of actions (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Whereas the generalization of these findings to humans is controversial (Dinstein, Thomas, Behrmann, & Heeger, 2008), neuroimaging studies (e.g., Etzel et al., 2008; Vogt et al., 2008; Buccino et al., 2004) have found evidence for the HMS that is thought to include the inferior frontal gyrus/frontal operculum, premotor cortex, and inferior parietal lobule (Rizzolatti, Fogassi, & Gallese, 2001). Importantly for the current results, primary motor cortex has also been shown to have similar mirror-like properties in nonhuman primates (Dushanova & Donoghue, 2010).

Because we did not directly compare action performance with action observation, we cannot directly comment on the HMS recruitment as a result of this study. We can, however, comment on regions that are recruited during action observation, one important component of the HMS. In the current study viewing actions, relative to fixation, recruited

the premotor cortex and inferior parietal lobule (although not the inferior frontal gyrus/frontal operculum) regardless of training experience, suggesting a role for action observation in this region. In contrast, active experience, compared with passive learning, resulted in the recruitment of the primary motor cortex upon subsequent perception of actions. This finding is potentially in contrast to M1 activity during action observation shown in the nonhuman primate (Dushanova & Donoghue, 2010) and serves to accentuate the different processing in the species studied. This conclusion should be tempered by the fact that with fMRI we are comparing relative differences between active and passive learning, and therefore, this region may still activate to some degree during both conditions. Given this, the current results still suggest that motor reactivation is more than just a simple mirroring effect that occurs when an action is being perceived because the activation primary motor region is modulated by previous active motor experience.

Functional Connectivity Is Affected by Learning Condition

Active versus passive learning also differed in subsequent functional connectivity between motor and visual processing regions during the processing of audiovisual stimuli. Butler et al. (2011) showed a similar finding in which a primary motor seed region showed a greater correlation with multiple visual processing regions including the LOC after active learning. In the current study, however, bilateral LOC seed regions showed stronger correlations with a right-lateralized premotor region. The difference in visual motor connectivity between these studies could be a result of the complexity and uniqueness of actions learned during training. In the current study, relatively more complex and unique actions were associated with each object, whereas in Butler et al. (2011), the same single grasping–reaching–pressing action was associated with each object. Therefore, the correlated premotor region, a higher-order motor processing region, may be utilized during learning and show modulation in subsequent visual motor connectivity. This increased coherence between activity in object-selective visual regions and higher-order motor processing regions suggests that active learning impacts not only subsequent reactivation but also subsequent connectivity. Furthermore, motor region reactivation recruited after active learning may be a result of strengthening of connections in a circuit linking sensory and motor processing.

Conclusion

In conclusion, learning audiovisual associations through self-generated actions resulted in faster and better-retained associative learning, motor/haptic system reactivation to audiovisual associations and enhanced functional connectivity between object-selective visual and premotor regions. Therefore, the current study extends previous

findings focusing on unisensory processing and supports theories of perceptual/motor reactivation by showing that active motor learning of sound producing objects impacts both audiovisual associative perception and recognition at behavioral and neural levels. Finally, the results of the current study clarify and extend our own previous work (Butler et al., 2011) by providing several novel findings and highlighting the task-based nature of motor reactivation and retrieval after active learning.

Acknowledgments

This research was partially supported by the Indiana METACyt Initiative of Indiana University, funded in part through a major grant from the Lilly Endowment, Inc. This research was also supported in part by the Faculty Research Support Program through the IU Bloomington Office of the Vice President of Research. We thank Ben Pruce and Chris Chung for their assistance with data collection and Shelley Swain with her help in stimuli creation. Also, we would like to especially thank Tom James for his help with both stimuli creation and his input on the experimental design.

Reprint requests should be sent to Andrew J. Butler, Butler University Psychology Department, 4600 Sunset Ave., Indianapolis, IN 46208, or via e-mail: Butler7@indiana.edu.

REFERENCES

- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–609.
- Bolognini, N., Frassinetti, F., Serino, A., & Ladavas, E. (2005). “Acoustical vision” of below threshold stimuli: Interaction among spatially converging audiovisual inputs. *Experimental Brain Research*, 160, 273–282.
- Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., Freund, H. J., et al. (2004). Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. *Neuron*, 42, 323–334.
- Butler, A. J., James, T. W., & Harman James, K. (2011). Enhanced multisensory integration and motor reactivation after active motor learning of audiovisual associations. *Journal of Cognitive Neuroscience*, 23, 3515–3528.
- Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage*, 50, 1148–1167.
- Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, 12, 478–484.
- Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33, 37–43.
- De Lucia, M., Camen, C., Clarke, S., & Murray, M. M. (2009). The role of actions in auditory object discrimination. *Neuroimage*, 48, 475–485.
- Dinstein, I., Thomas, C., Behrmann, M., & Heeger, D. J. (2008). A mirror up to nature. *Current Biology*, 18, 13–18.
- Dushanova, J., & Donoghue, J. (2010). Neurons in primary motor cortex engaged during action observation. *European Journal of Neuroscience*, 31, 386–398.
- Etzel, J. A., Gazzola, V., & Keysers, C. (2008). Testing simulation theory with cross-modal multivariate classification of fMRI data. *PLoS*, 3, 1–6.
- Fredembach, B., de Boisferon, A., & Gentaz, E. (2009). Learning of arbitrary association between visual and auditory stimuli

- in adults: The “bond effect” of haptic exploration. *PLoS*, *4*, e4844.
- Fuster, J. M. (2009). Cortex and memory: Emergence of a new paradigm. *Journal of Cognitive Neuroscience*, *21*, 2047–2072.
- Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 67–82). Hillsdale, NJ: Erlbaum.
- Gottfried, J. A., Smith, A. P. R., Rugg, M. D., & Dolan, R. J. (2004). Remembrance of odors past: Human olfactory cortex in cross-modal recognition memory. *Neuron*, *42*, 687–695.
- Grezes, J., & Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia*, *40*, 212–222.
- Grezes, J., Tucker, M., Armony, J., Ellis, R., & Passingham, R. E. (2003). Objects automatically potentiate action: An fMRI study of implicit processing. *European Journal of Neuroscience*, *17*, 2735–2740.
- Harman, K. L., Humphrey, G. K., & Goodale, M. A. (1999). Active manual control of object views facilitates visual recognition. *Current Biology*, *9*, 1315–1318.
- Hornberger, M., Rugg, M. D., & Henson, R. N. A. (2006). fMRI correlates of retrieval orientation. *Neuropsychologia*, *44*, 1425–1436.
- Imamizu, H., Kuroda, T., Miyauchi, S., Yoshioka, T., & Kawato, M. (2003). Modular organization of internal models of tools in the human cerebellum. *Proceedings of the National Academy of Sciences, U.S.A.*, *100*, 5461–5466.
- James, K. H., & Atwood, T. P. (2009). The role of sensorimotor learning in the perception of letter-like forms: Tracking the causes of neural specialization for letters. *Cognitive Neuropsychology*, *26*, 91–110.
- James, K. H., & Gauthier, I. (2006). Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia*, *44*, 2937–2949.
- James, K. H., Humphrey, G. K., & Goodale, M. A. (2001). Manipulating and recognizing virtual objects: Where the action is. *Canadian Journal of Experimental Psychology*, *55*, 111–120.
- James, K. H., Humphrey, G. K., Vilis, T., Baddour, R., Corrie, B., & Goodale, M. A. (2002). Learning three-dimensional object structure: A virtual reality study. *Behavioral Research Methods, Instruments and Computers*, *34*, 383–390.
- James, K. H., & Mauoene, J. (2009). Auditory verb perception recruits motor systems in the developing brain: An fMRI investigation. *Developmental Science*, *12*, F26–F34.
- James, K. H., & Swain, S. N. (2011). Only self-generated actions create sensori-motor systems in the developing brain. *Developmental Science*, *13*, 279–288.
- James, T. W., Culham, J., Humphrey, G. K., Milner, D. A., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: A fMRI study. *Brain*, *126*, 2463–2475.
- James, T. W., & Gauthier, I. (2003). Auditory and action semantic features activate sensory-specific perceptual brain regions. *Current Biology*, *13*, 1792–1796.
- Kahn, I., Davachi, L., & Wagner, A. D. (2004). Functional-neuroanatomic correlates of recollection: Implications for models of recognition memory. *The Journal of Neuroscience*, *24*, 4172–4180.
- Longcamp, M., Anton, J.-L., Roth, M., & Velay, J.-L. (2005). Premotor activations in response to visually presented single letters depend on the hand used to write: A study on left-handers. *Neuropsychologia*, *43*, 1801–1809.
- Masumoto, K., Yamaguchi, M., Sutani, K., Tsuneto, S., Fujita, A., & Tonoike, M. (2006). Reactivation of physical motor information in the memory of action events. *Brain Research*, *1101*, 102–109.
- Murray, M. M., Foxe, J. J., & Wylie, G. R. (2005). The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage*, *27*, 473–478.
- Murray, M. M., & Sperdin, H. F. (2010). Single-trial multisensory learning and memory retrieval. In M. J. Naumer & J. Kaiser (Eds.), *Multisensory object perception in the primate brain* (pp. 191–208). New York: Springer Science + Business Media, LLC.
- Mutschler, I., Schulze-Bonhage, A., Glauche, V., Demandt, E., Speck, O., & Ball, T. (2007). A rapid sound-action association effect in human insular cortex. *PLoS*, *2*, 1–9.
- Nilsson, L.-G., Nyberg, L., Klingberg, T., Aberg, C., Persson, J., & Roland, P. E. (2000). Activity in motor areas while remembering action events. *NeuroReport*, *11*, 2199–2201.
- Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences*, *9*, 11120–11124.
- Nyberg, L., Petersson, K. M., Nilsson, L.-G., Sandblom, J., Aberg, C., & Ingvar, M. (2001). Reactivation of motor brain areas during explicit memory for actions. *Neuroimage*, *14*, 521–528.
- Pulvermuller, F., Harle, M., & Hummel, F. (2001). Walking or talking: Behavioral and neurophysiological correlates of action verb processing. *Brain and Language*, *78*, 143–168.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding of imitation of action. *Nature Reviews Neuroscience*, *2*, 661–670.
- Sasaoka, T., Asakura, N., & Kawahara, T. (2010). Effect of active exploration of 3-D object views on the view-matching process in object recognition. *Perception*, *39*, 289–308.
- Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional localizers. *Neuroimage*, *30*, 1088–1096.
- Senkfor, A. J., Petten, C. V., & Kutas, M. (2002). Episodic action memory for real objects: An ERP investigation with perform, watch, and imagine action encoding tasks versus a non-action encoding task. *Journal of Cognitive Neuroscience*, *14*, 402–419.
- Slotnick, S. D., & Schacter, D. L. (2006). The nature of memory related activity in early visual areas. *Neuropsychologia*, *44*, 2874–2886.
- Stevenson, R. A., Altieri, N. A., Kim, S., Pisoni, D. B., & James, T. W. (2010). Neural processing of asynchronous audiovisual speech perception. *Neuroimage*, *49*, 3308–3318.
- Stevenson, R. A., Geoghegan, M. L., & James, T. W. (2007). Superadditive BOLD response in superior temporal sulcus with threshold non-speech objects. *Experimental Brain Research*, *179*, 85–95.
- Stevenson, R. A., & James, T. W. (2009). Neuronal convergence and inverse effectiveness with audiovisual integration of speech and tools in human superior temporal sulcus: Evidence from BOLD fMRI. *Neuroimage*, *44*, 1210–1223.
- Stock, O., Roder, B., Burke, M., Bien, S., & Rosler, F. (2009). Cortical activation patterns during long-term memory retrieval of visually or haptically encoded objects and locations. *Journal of Cognitive Neuroscience*, *21*, 58–82.
- Talairach, J., & Tournoux, P. (1988). *A co-planar stereotactic atlas of the human brain: 3-Dimensional proportional system: An approach to cerebral mapping* (M. Rayport, Trans.). New York: Thieme.

- Vaidya, C. J., Zhao, M., Desmond, J. E., & Gabrieli, J. D. E. (2002). Evidence for cortical encoding specificity in episodic memory: Memory-induced re-activation of picture processing areas. *Neuropsychologia*, *40*, 2136–2143.
- Vilberg, K. L., & Rugg, M. D. (2007). Dissociation of the neural correlates of recognition memory according to familiarity, recollection, and amount of recollected information. *Neuropsychologia*, *45*, 2216–2225.
- Vogt, S., Buccino, G., Wohlschlager, A. M., Canessa, N., Shah, N. J., Zilles, K., et al. (2008). Prefrontal involvement in imitation learning of hand actions: Effects of practice and expertise. *Neuroimage*, *37*, 1371–1383.
- Weisberg, J., van Turenout, M., & Martin, A. (2007). A neural system for learning about object function. *Cerebral Cortex*, *17*, 513–521.
- Wheeler, M. E., Peterson, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences*, *97*, 11125–11129.
- Wheeler, M. E., Shulman, G. L., Buckner, R. L., Miezin, F. M., Velanova, K., & Peterson, S. E. (2006). Evidence for separate perceptual reactivation and search processes during remembering. *Cerebral Cortex*, *16*, 949–959.