Only self-generated actions create sensori-motor systems in the developing brain

Karin Harman James[1,2,3] and **Shelley N. Swain**[1]

[1]Department of Psychological and Brain Sciences, Indiana University, USA

[2]Program in Cognitive Science, Indiana University, USA

[3]Program in Neuroscience, Indiana University, USA

Abstract

Previous research shows that sensory and motor systems interact during perception, but how these connections among systems are created during development is unknown. The current work exposes young children to novel '*verbs*' and objects through either (a) actively exploring the objects or (b) by seeing an experimenter interact with the objects. Results demonstrate that the motor system is recruited during auditory perception only after learning involved self-generated interactions with objects. Action observation itself led to above-baseline activation in one motor region during visual perception, but was still significantly less active than after self-generated action. Therefore, in the developing brain, associations are built upon real-world interactions of body and environment, leading to sensori-motor representations of both objects and words.

Introduction

Experience results in the co-activation of sensory and motor systems in the brain. Recent research indicates that motor systems are active during visual object perception if the object is associated with a history of action (e.g. Chao & Martin, 2000). Motor systems are also active in the processing of some verbs, another kind of stimulus that is associated with action (e.g. Pulvermüller, Harle & Hummel, 2001; Hauk, Johnsrude & Pulvermüller, 2004). Other evidence indicates that it is not just verbs that activate the motor system, but any word that has a history of association with an action – including nouns (Saccuman, Cappa, Bates, Arevalo, Della Rosa, Danna & Perani, 2006; Arevalo, Perani, Cappa, Butler, Bates & Dronkers, 2007). These findings are changing contemporary understanding of the multimodal and sensori-*motor* nature of the processes that underlie perception and cognition (Barsalou, Simmons, Barbey & Wilson, 2003). What we do not know, however, is how these sensori-motor connections are formed in the first place.

Recent research shows that verb processing activates motor systems in the *developing* brain (James & Maouene, 2009). In this work, verbs that referred to hand movements or leg

movements activated the motor cortex significantly more than adjectives. Importantly, the activation was effector specific – verbs associated with hand actions recruited different regions of the motor cortex than verbs associated with leg movements. Thus, in 4–6-year-old children verb perception was associated with effector-specific motor system activation. This suggests that links among sensory and motor systems are being created early in development when experience with the world and knowledge of verbs is still changing rapidly.

However, the type of experience that might be required for motor system recruitment during perception is the subject of some controversy. Classic developmental theories point to self-generated action, to doing (e.g. Piaget, 1952), but there is little relevant empirical evidence. More recent research suggests that in adults simply observing another performing an action on an object will result in motor system activation (e.g. Rizzolatti & Craighero, 2004). This suggests that merely watching others perform labeled actions or actions on objects might be sufficient to lead to motor system activation when subsequently perceiving the verb or the object. The present research provides new evidence relevant to these competing hypotheses and does so in young children, the period in which the connections so evident in adults are presumably being built. We show for the first time the kinds of experiences that are necessary for motor systems to be recruited during perception.

Methods

Participants

Participants were 13 children between the ages of 5 years 5 months and 6 years 9 months (11 males). All children had normal or corrected to normal vision and had normal hearing as reported by parents. Participants were not on any medication and had no history of neurological compromise. All were delivered at term without a record of birth trauma. Parents reported right-hand dominance for all participants.

General procedure

Participants underwent both a training session and a test session. The training session was performed outside of the fMRI environment. There were two within-participant conditions: one in which the participants performed an action on an object themselves (active condition), and another in which the experimenter performed the action and the participant observed her (passive condition). The actions were performed on novel, 3-D objects (see Figure 1), resulting in each case in associations among an action, a novel object and a verb. After training, participants were tested on their knowledge of the novel action labels to ensure that they had learned the labels. Stimulus exposure in the active and passive conditions was equated for each participant. Subsequent to this training, an fMRI session was performed to test (1) whether or not the learned actions would activate the motor cortex when the novel action labels were heard (auditory perception) and / or when the novel objects were seen (visual perception); and (2) in each perception condition, whether motor system activation would occur after self-generated action (active condition) and / or after observing the action of another (passive condition). As control conditions, novel verbs and novel objects that were not experienced in the training session were also presented to the participants.

Training stimuli

Novel action labels followed standard verb morphology and were all of equal length and complexity. Participants learned 10 novel verb labels (yocking, wilping, tifing, sprocking, ratching, quaning, panking, nooping, manuing, leaming), half during active interaction with objects and half during passive observation of actions. The objects that were acted upon were novel, three-dimensional plastic objects, painted in monochromatic primary colors. The objects were approximately $12 \times 8 \times 6$ cm and weighed approximately 115 g. Each object was constructed from 2–3 primary shapes (see Figure 1). The objects could be acted upon, which could change their shape, and each action was unique to each object. For example, object A was associated with 'sprocking' – pulling out a retractable cord from its center. When the objects were not acted upon, the action was not afforded by their appearance alone – that is, it was not obvious how to interact with the object upon visual perception.

Training procedure

The participant and an experimenter sat across from each other at a table. Each had an object set (five objects) laid out randomly in a straight horizontal line directly in front of them. All objects were in full view of the participant. Object sets were counterbalanced across participants. The training procedure was structured like a game, engaging the participants, resulting in all participants completing the training.

A second experimenter acted as the referee and directed the session. The participant was told that they would be playing with some toys that have a specific action and that action has a special name. The referee demonstrated the action associated with each toy and said the name of the action. In the passive, or observation condition, the experimenter's five action names were placed in a bag and the experimenter drew a name from the bag, said it, then chose the appropriate object and performed the appropriate action five times. While performing the action on the object the experimenter continued to say the action name and encouraged the participant to also say the action name (e.g. Look, I'm yocking it, wow cool, I'm yocking it. What am I doing?). This was repeated until all action names were drawn. In the active, self-generated action condition, the procedure was the same, with the child randomly choosing an object and action, but the referee named the action at first, until the child was able to name the action on their own. In both the active and passive training, the children were encouraged to say the action label each time the action was performed. If they did not, the experimenter reminded them to do so, resulting in all children saying all labels each time an action was performed. The referee also aided the participant in choosing the correct object and performing the correct action if it was required. After the participant drew all of the action names they were rewarded with a sticker and then it was the experimenter's turn again (passive condition). This process was repeated five times in total, resulting in the experimenter and participant interacting with their objects 25 times (5×5 active objects and 5×5 passive objects: 50 total exposures). Subsequent to the training procedure, the experimenter randomly selected each object, showed it to the child, and asked them to say the name of the label associated with each of the 10 objects. All children were able to name all the actions associated with the objects on the first trial, ensuring that they had learned all

labels, both those learned actively and passively. Duration of session was approximately 20–30 minutes.

fMRI test stimuli

Auditory stimuli were the action labels that were learned from the training session as well as five new words that followed the same morphological rules as the trained verbs. These words were read by a female voice and were presented to the participants through headphones in the fMRI facility. Visual stimuli were photographs of the learned objects as well as five photos of similar, unlearned objects. The photos depicted the objects from a variety of planar (axis of elongation was 0 degrees from the observer) and ¾ (axis of elongation was 45 degrees from the observer) viewpoints. The main features of the objects could be seen in every photo.

Testing procedure

After screening and informed consent given by the parent, all participants were acclimated to the MRI environment by watching a cartoon in an MRI 'simulator'. The simulator is the same dimension as the actual MRI and the sound of the actual MRI environment is played in the simulator environment. This allowed the children to become comfortable in the environmental set-up before entering the actual MRI environment. After the participant felt comfortable in this environment, and if the parent was comfortable with the participant continuing, they were introduced to the actual MRI environment.

Following instructions, a preliminary high-resolution anatomical scan was administered while the participant watched a cartoon. Following this scan, the functional scanning occurred. Auditory and visual stimuli were separated into different runs, and each run consisted of six blocks of stimuli (two 'active', two 'passive', two 'novel'). Blocks were 18–20 seconds long and there were 10 second intervals between blocks. Each run began with a 20 second rest period and ended with a 10 second rest period. This resulted in runs that were just under 3.5 minutes long. Two to four runs were administered depending on the comfort of the child. During auditory presentation runs, participants were required simply to passively listen to the stimuli. The action words were read at a rate of approximately 1.25 seconds per word, and new, actively learned and passively learned action words were separated into separate blocks. During the visual perception runs, participants were required to passively view the stimuli that were presented for 1 second each. New, actively learned and passively learned objects were presented in separate blocks. Neural activation, measured by the BOLD (Blood Oxygen Level Dependent) signal in the entire brain, was then recorded during exposure to the stimuli. Imaging sessions took approximately 15 minutes in total.

After the functional scans, the participant was removed from the environment, debriefed, and rewarded for their time.

fMRI data acquisition

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. Whole-brain axial images were acquired using an echo-planar

technique (TE = 30 ms, TR = 2000 ms, flip angle = 90°) for BOLD based imaging. The field of view was 22 × 22 × 9.9 cm, with an in plane resolution of 64 × 64 pixels and 33 slices per volume that were 4 mm thick with a 0 mm gap among them. The resulting voxel size was 3.0 mm × 3.0 mm × 4.0 mm. Functional data underwent slice time correction, 3-D motion correction, linear trend removal, and Gaussian spatial blurring (FWHM 6 mm) using the analysis tools in Brain Voyager™. Individual functional volumes were co-registered to anatomical volumes with an intensity matching, rigid-body transformation algorithm. Voxel size of the functional volumes was standardized at 1 mm × 1 mm × 1 mm using trilinear interpolation. High-resolution T1-weighted anatomical volumes were acquired prior to functional imaging using a 3-D Turbo-flash acquisition (resolution: 1.25 mm × 0.62 × 0.62, 128 volumes).

Data analysis procedures

Whole-brain Group contrasts were performed on the resultant data. The functional data were analyzed with a random effects general linear model (GLM) using Brain Voyager's™ multi-subject GLM procedure. The GLM analysis allows for the correlation of predictor variables or functions with the recorded activation data (criterion variables) across scans. The predictor functions were based on the blocked stimulus presentation paradigm of the particular run being analyzed and represent an estimate of the predicted hemodynamic response during that run. Any functional data that exceeded 5 mm of motion on any axis were excluded from the analyses. This criterion resulted in excluding two blocks of data from one participant and one block of data from three participants. Exclusion of these data does not significantly alter the power of the present analyses. Data were transformed into a common stereotactic space (e.g. Talairach & Tournoux, 1988) for group-based statistical analyses. Direct contrasts of BOLD activation were performed on the group between active action labels and passive action labels that were learned (new action words were used as a baseline). In addition, contrasts between activation during perception of objects that were learned actively vs. passively were also analyzed.

Contrasts in the group statistical parametric maps (SPMs) were considered above threshold if they met the following criteria in our random effects analysis: (1) significant at $p < .001$, uncorrected, with a cluster threshold of 270 contiguous 1 mm isometric voxels; (2) peak statistical probability within a cluster at least $p < .0001$, uncorrected.

Results and discussion

Auditory verb perception

Results indicated that when comparing participants' BOLD responses to learned vs. unlearned novel verbs, motor systems were activated by learned verbs only when the verbs were learned through active interaction with the objects (see Figure 2A) (see Table 1 for co-ordinates, cluster sizes and significance). Hearing actively learned novel action labels resulted in significantly greater activation than passively learned labels or unlearned labels in the right superior frontal gyrus (Brodmann's area 6); the left middle frontal gyrus (6); and the inferior parietal lobule, bilaterally (Brodmann's area 40). The motor system in the right frontal cortex was recruited only after active experience with objects while learning verb

names. Interestingly, this activation pattern was also observed in the inferior parietal lobe, bilaterally, a region associated with grasping in humans and monkeys (e.g. Culham & Valyear, 2006) but also, and perhaps more relevantly, with tool use (Johnson-Frey, 2004). If our novel objects are represented similarly as known tools after active interaction, then activation in this area is not surprising. However, finding this activation only during auditory perception and not visual perception (see below) is somewhat more novel. Possibly the role of the inferior parietal lobule is not in tool representation as much as it is involved in action representation – being close to parietal regions associated with grasping. Action labels (the novel verbs) may recruit regions that are associated with the actual actions, whereas seeing the object may not activate the actual action patterns associated with interaction, but rather the sensori-motor representation associated with the frontal cortices.

Visual object perception

Seeing objects that were actively interacted with during training resulted in greater BOLD signal change than seeing objects that were passively observed during training in the left precentral gyrus (Brodmann's area 4). No other brain regions differed in their neural activation to the two types of learned objects. Thus, when comparing responses to learned vs. unlearned novel objects, motor systems were recruited during learned object perception more than during unlearned object perception, but again, more so if the objects were learned through active interaction (Figure 2B).

In the left frontal gyri, both active learning and passive learning were above baseline activation (see Figure 2A, upper right side and Figure 2B), although still significantly different from one another. During auditory perception this occurred in the middle frontal gyrus, and during visual perception in the left precentral gyrus. This finding suggests that both action production and action observation recruit these regions, a finding that supports numerous studies in adult humans and non-human primates (for review see Rizzolatti & Craighero, 2004). However, we show that active learning does activate these regions to a greater degree than does passive learning – shown by the significant difference produced by the direct contrast between active and passive experience. In addition, to confirm our observations, when we performed a direct contrast between passive learning and baseline during auditory perception, no brain regions were significantly active (this null result is not shown), but during visual processing the precentral gyrus was active. Therefore, the left precentral gyrus is active during visual perception of objects that were learned through active interaction and through passive observation, but more so during perception of objects that were previously learned through self-generated actions.

Therefore, self-generated actions were required for the emergence of motor system recruitment during auditory processing in the developing brain, but both active and passive learning recruited left motor regions during visual perception. Numerous theories have suggested that action and perception, when coupled, form representations that may be accessed by perception alone (e.g. Prinz, 1997) – that these representations contain the motor programs associated with the percept. In addition, performed actions will activate visual cortices without concurrent visual stimulation just as perception can activate motor systems without concurrent movement (e.g. James & Gauthier, 2006). The frontal system

codes information that associates previously performed actions with present perceptions, and is therefore recruited to a significantly lesser extent during perception after action observation. This finding appears, on the surface, to stand at odds with work showing frontal activation during action observation (e.g. Rizzolatti & Craighero, 2004), but given the above baseline activation for action observation in the left precentral gyrus, it actually supports this and other findings on static object perception (e.g. Chao & Martin, 2000) as well as perception of actions.

It is possible that actively interacting with objects allows for the participants to 'imagine' the actions upon subsequent auditory or visual presentation of the objects, resulting in differences between active and passive experience. That is, covert enactment of the motor patterns associated with the actively learned objects may have recruited the brain regions in this study (e.g. Jeannerod, 2001). This would contrast somewhat with our interpretation that the sensori-motor information gets associated through the learning experience, resulting in an activation pattern that accesses both sensory and motor information directly. The current work cannot distinguish between these two alternatives, as we cannot ascertain whether activation seen here is due to imagining actions or automatic access to learned actions. Either way, however, active experience changes how the brain processes subsequent auditory or visual information. The level at which this experience affects subsequent perception, be it from directly accessing prior motor activity or through allowing actions to be imagined, is an important question for future work that can distinguish timing information in neural processing – an obvious shortcoming of fMRI blocked designs (e.g. Hauk, Shtyrov & Pulvermüller, 2008).

In the adult brain, we know that action words, and in some cases object perception, recruits motor systems. Here we show that in order for this adult-like sensori-motor response to occur, children need to actively interact with objects in the environment. Therefore, we provide initial evidence for the types of interactions that produce adult-like neural responses. Providing such developmental information is important for understanding the cascading effects that certain experiences have on neural response patterns and, therefore, on human cognition. Obviously children do not have the rich semantic representations for words and objects that adults acquire during their additional years of experience, and thus we may expect their representations to be more sensori-motor in nature than those of adults. Our knowledge regarding differences in the degree to which representations are predominantly sensori-motor and / or semantic / conceptual throughout development would be enhanced by future work on this topic.

This work allows us to come closer to understanding the role of the frontal and parietal systems for object and action-word processing. Based on our present results, we propose that one function of the motor association areas is to associate past experience with present perception, but in a fairly specific manner – associating a history of self-generated actions with perception. That is, at least in the developing brain, perception and action become strongly linked as a result of self-generated action: in general, experience must be sensori-motor and not sensation-of-motor.

Acknowledgments

References

Arevalo A, Perani D, Cappa SF, Butler A, Bates E, Dronkers N. Action and object processing in aphasia: from nouns and verbs to the effect of manipulability. Brain and Language. 2007; 100:79–94. [PubMed: 16949143]

Barsalou LW, Simmons WK, Barbey A, Wilson CD. Grounding conceptual knowledge in modality-specific systems. Trends in Cognitive Sciences. 2003; 7:84–91. [PubMed: 12584027]

Chao LL, Martin A. Representation of man-made objects in the dorsal stream. NeuroImage. 2000; 12:478–484. [PubMed: 10988041]

Culham JC, Valyear KF. Human parietal cortex in action. Current Opinion in Neurobiology. 2006; 16:205–212. [PubMed: 16563735]

Hauk O, Johnsrude I, Pulvermüller F. Somato-topic representation of action words in human motor and premotor cortex. Neuron. 2004; 41:301–307. [PubMed: 14741110]

Hauk O, Shtyrov Y, Pulvermüller F. The time course of action and action-word comprehension in the human brain as revealed by neurophysiology. Journal of Physiology, Paris. 2008; 1:50–58.

James KH, Gauthier I. Letter processing automatically recruits a sensory-motor brain network. Neuropsychologia. 2006; 44:2937–2949. [PubMed: 16920164]

James KH, Maouene J. Auditory verb perception recruits motor systems in the developing brain: an fMRI investigation. Developmental Science. 2009; 12:F26–F34. [PubMed: 19840036]

Jeannerod M. Neural simulation of action: a unifying mechanism for motor cognition. NeuroImage. 2001; 14:S103–S109. [PubMed: 11373140]

Johnson-Frey SH. The neural basis of complex tool use in humans. Trends in Cognitive Sciences. 2004; 8:227–237.

Piaget, J. The origins of intelligence in children. New York: International University Press; 1952.

Prinz W. Perception and action planning. European Journal of Psychology. 1997; 9:129–154.

Pulvermüller F, Harle M, Hummel F. Walking or talking: behavioral and neurophysiological correlates of action verb processing. Brain and Language. 2001; 78:143–168. [PubMed: 11500067]

Rizzolatti G, Craighero L. The mirror-neuron system. Annual Review of Neuroscience. 2004; 27:169–192.

Saccuman MC, Cappa SF, Bates EA, Arevalo A, Della Rosa P, Danna M, Perani D. The impact of semantic reference on word class: an fMRI study of action and object naming. NeuroImage. 2006; 32:1865–1878. [PubMed: 16766208]

Talairach, J.; Tournoux, P. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system – an approach to cerebral imaging. New York: Thieme Medical Publishers; 1988.

Context of observation:

A Ostensive communication:

Child

E2 (Max)   E1

B Incidental observation:

45 cm   Child

E1

E2   E1
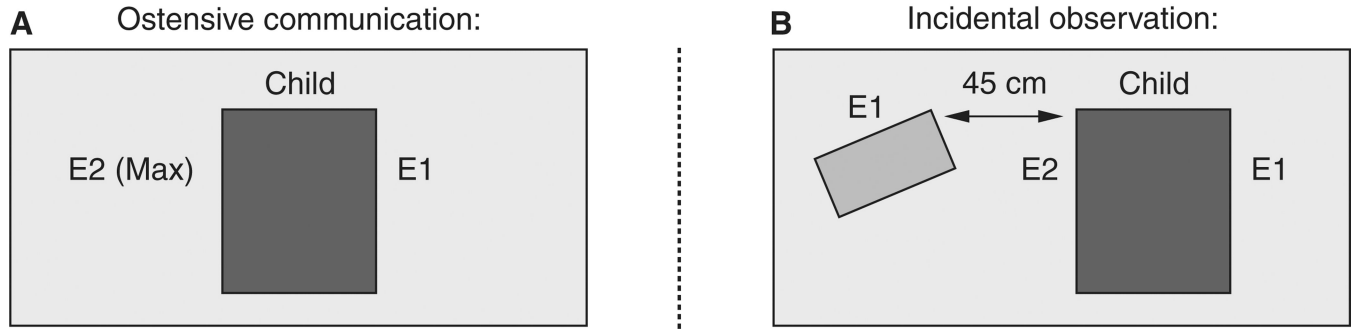
Figure 1.
(A) Examples of 3-D plastic objects used during training. (B) Examples of actions and action labels associated with objects during training.
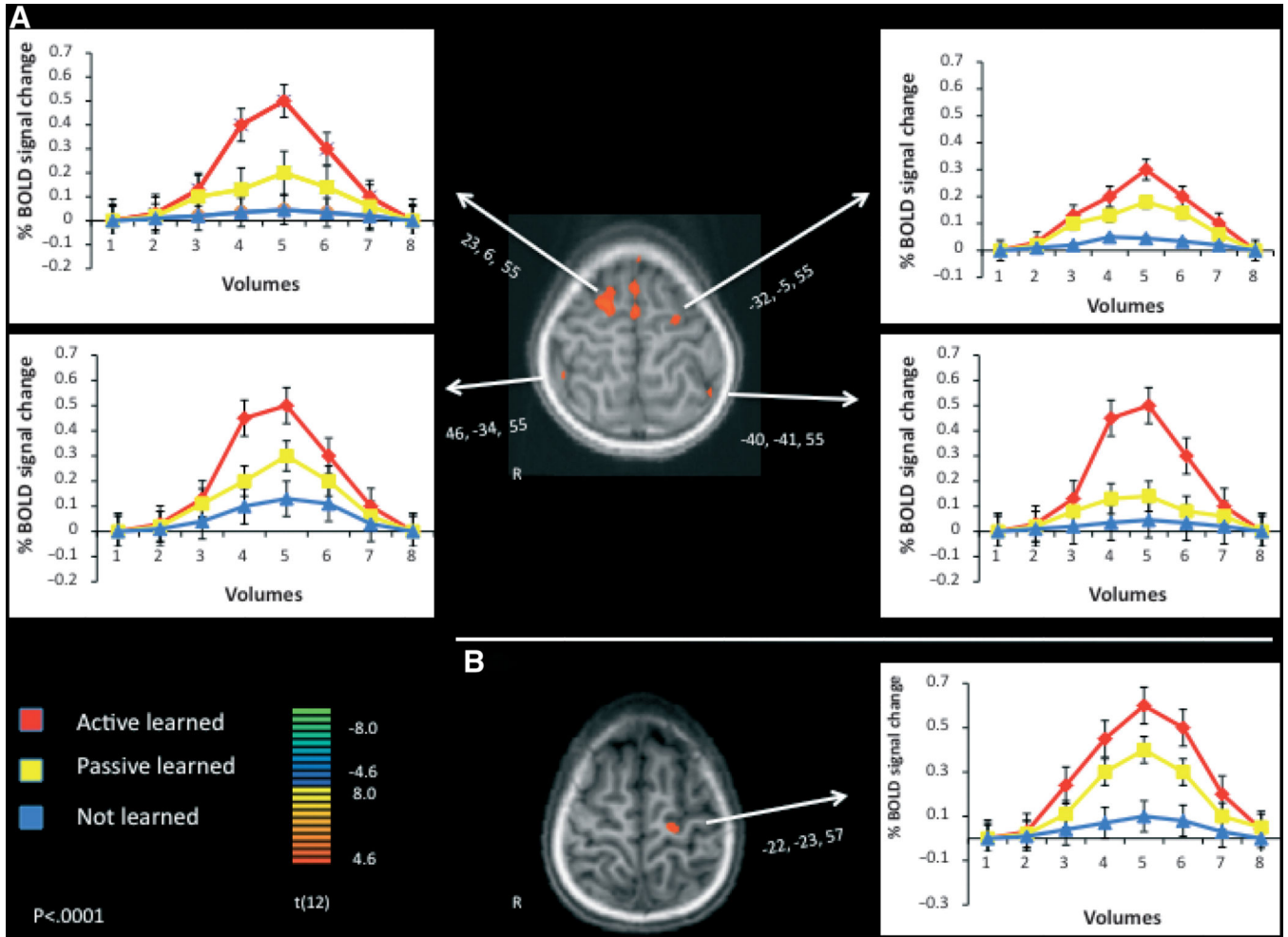
Figure 2.
Averaged BOLD percent signal change comparing active learning (positive values) and passive learning. Clusters were considered significant if they exceeded 10 contiguous voxels at a threshold of p < .0001, uncorrected. (A) Hearing actively learned novel action labels resulted in significantly greater activation than passively learned labels or unlearned labels in the right superior frontal gyrus (Brodmann's area 6); the left middle frontal gyrus (6); and the inferior parietal lobule, bilaterally (40). (B) Seeing objects that were actively interacted with during training resulted in greater % signal change than seeing objects that were passively observed during training or seeing unlearned objects, in the left precentral gyrus (4). No motor responses were given during the fMRI sessions. Coordinates listed are from Talairach and Tourneaux (1988). Error bars depict between-subjects standard error of the mean.

Table 1

Anatomical regions from whole-brain contrasts, peak Talairach coordinates, cluster size and peak *t*-value

Regions	Peak Talairach coordinates (x, y, z)	Cluster size	Peak *t*(12)	Significance (*p* <)
Right superior frontal gyrus	23,6,55	961	6.1	.00001
Right inferior parietal lobule	46,−34,55	275	4.7	.001
Left middle frontal gyrus	−32,−5,55	554	6.0	.0001
Left inferior parietal lobule	−40,−41,55	310	4.9	.001
Left precentral gyrus	−22,−23,57	430	5.8	.0001