



Published in final edited form as:

*Cogn Brain Behav.* 2011 ; 15(4): 485–503.

## SELF-GENERATED ACTIONS DURING LEARNING OBJECTS AND SOUNDS CREATE SENSORI-MOTOR SYSTEMS IN THE DEVELOPING BRAIN

Karin Harman JAMES<sup>1,2,3,\*</sup> and Paroma BOSE<sup>1</sup>

<sup>1</sup>Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

<sup>2</sup>Program in Cognitive Science, Indiana University, Bloomington, IN, USA

<sup>3</sup>Program in Neuroscience, Indiana University, Bloomington, IN, USA

### Abstract

Previous research shows that sensory and motor systems interact during verb perception, and that these interactions are formed through self-generated actions that refer to verb labels during development. Here we expand on these findings by investigating whether self-generated actions lead to sensori-motor interaction during sound perception and visual perception. The current research exposes young children to novel sounds that are produced by object movement through either a) actively exploring the objects and producing the sounds or b) by seeing and hearing an experimenter interact with the objects. Results demonstrate that the motor system was recruited during auditory perception only after learning involved self-generated interactions with objects. Interestingly, visual association regions were also active during both sound perception and visual perception after active exploratory learning, but not after passive observation. 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 sounds.

### Keywords

fMRI; embodiment; brain development

---

Because individuals are situated in an environment that they actively explore, vision and motor experience are associated during learning about objects. 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; Grezes & Decety, 2002; Longcamp, Anton, Roth & Velay, 2005; James & Gauthier, 2006; Weisberg, van Turennout & Martin, 2007). Thus, embodied experiences result in the co-activation of sensory and motor systems in the brain. Although controversial, mounting research suggests that observation of action in humans is very different from actual self-generated action in terms of brain responses in children (James & Swain, 2011) and in adults (Butler & James, 2011, Turella, Pierno,

Tubaldi, & Castiello, 2009; Dinstein, Hasson, Rubin, & Heeger, 2007; see also Dinstein, Gardner, Jazayeri, & Heeger, 2008), although others show an overlap of responses during observed and executed actions (Iacoboni, 2005; Rizzolatti & Craighero, 2004). Motor systems are also active in the processing of some verbs, another kind of stimulus that is associated with action (eg. Pulvermüller, Harle, & Hummel, 2001; Hauk, Johnsrude & Pulvermüller, 2004; James & Maouene, 2009). 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 et al., 2006; Arevalo et al., 2007). These findings are changing contemporary understanding of the multi-modal and sensori-*motor* nature of the processes that underlie perception and cognition (Barsalou, et al., 2003).

In the developing system, it stands to reason that action is crucial for learning. This is not a new idea, being proposed by such leaders in our field as Piaget (1977) and Gibson (1969), and more recently has been elaborated upon by several developmental psychologists (Thelan & Smith, 1994; Bertenthal, & Campos, 1987; Bushnell & Boudreau, 1993; Needham et al., 2002). However, the mechanisms that support this important interaction are not well understood. By studying neural responses during learning with and without action, we can begin to understand *why* action is important to learning in many domains. Thus far, we have documented that action changes brain processing in children during cognitive operations such as verb processing (James & Maouene, 2009) object perception (James & Swain, 2011), and letter perception (James, 2010). However, only one study to date has directly compared active learning directly with passive observation of action in young children using fMRI (James & Swain, 2011). In that study, children were trained on learning novel verbs that were associated with actions on objects. The actions that were associated with the verbs were either learned actively, through the child producing the action and naming it, or passively, by observing an experimenter produce the action and name it. In both situations, the children learned the verbs and the actions, but when functional imaging was performed, only the active condition resulted in sensori-motor networks responding to the verb names. Passively learning the verbs while watching others did not recruit a motor network, whereas active learning did (James & Swain, 2011). From an embodied framework, it makes sense that self-generated actions that produced certain percepts - be they sound or image percepts – would result in sensory-motor associations in the brain. When learning about objects, simply watching another act on an object to produce images and sounds did not appear to result in motor reactivation in the brain (James & Swain, 2011), lending support to the idea that our own actions on things – our embodiment – is important for learning. The question remains however, whether this sensory-motor association is specific to word processing, or if any sound will produce such co-activation of these systems after self-generated learning.

As our primary form of communication, the understanding and production of speech is a behavior that occurs rapidly and seemingly effortlessly within a typical learning environment. Because spoken language is learned in young children in every culture and under almost all circumstances, it is thought by some to be a unique human behavior (eg. Chomsky & Ronat, 1998). In addition, language perception has special status in the brain. That is, humans have specialized systems that process language – presumably these systems have evolved to allow us to learn and understand language at an astonishing rate (eg.

MacWhinney, 1999). This is most apparent when young children learn their mother tongue, but continues through childhood, up to the point when learning language becomes more effortful (eg. Bialystok, 1991). It is possible that learning words that are associated with actions leads to motor system re-activation during subsequent perception, but that other sounds may not. This would support the idea that learning language is a special form of perception, crucial for human development. However, an alternate explanation is that systems in the brain become associated through any self-generated experience –language is not special in this regard. This latter hypothesis would predict that any sensory-motor association formed during learning would result in motor system re-activation during subsequent perception-be it visual, auditory, haptic, etc.

The current study addresses these competing hypotheses by requiring children to learn associations among actions and sounds that are not words. If language is special in producing auditory-motor associations in the brain, then motor re-activation should not be seen after learning in this case. However, if motor re-activation does emerge after this learning, then it is the association that results in re-activation and not language specifically. Results of our research suggest the latter – that any association during learning will result in sensory-motor associations forming in the neural systems that support these functions. Further, we test visual perception of actions after active vs passive learning for two reasons: a) to compare our findings here with our previous work on visual perception after active learning (James, 2010; Butler & James, 2011; James & Swain, 2011), and b) to see if viewing *videos* are different from viewing static images after learning. However, these finding are not stressed here, as our primary goal was to investigate sound perception.

## METHODS

### Participants

Participants were 18 children between the ages of 5 years 1 months and 7 years 1 month (9 males). All children had normal or corrected to normal vision and normal hearing as reported by parents. Participants were not on any medications and had no history of neurological compromise. All were delivered at term without a record of birth trauma. Parents reported right-hand dominance for 16 participants and left-hand dominance for 2 participants. Functional imaging data from four participants was not gathered due to discomfort in the MRI environment. Thus, data from 14 individuals is reported here, 7 males, ranging from 5 years 4 months to 7 years 1 month.

### 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 that produced a sound when a specific action was performed on them (see Figure 1). This resulted in learning an association among an unlabelled action, a novel object, and a sound. 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 sounds were heard (auditory perception), when the novel objects were seen (visual perception), and/or when the novel objects were heard and seen (auditory + 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 sounds and novel objects that were not experienced in the training session were also presented to the participants.

### Training Stimuli

The objects used had particular actions and sounds associated with them. The actions themselves could be performed at different time scales (faster, slower), and therefore the sound produced and the actions were somewhat reliant of the actor. However, we tried to equate the actions as much as possible (and therefore the resultant sounds they made). The actions and their associated sounds are presented in Table 1. Participants learned actions/sounds for 10 novel objects, 5 through active interaction with objects and 5 through passive observation of actions. The objects acted upon were novel, three-dimensional plastic objects, painted with monochromatic primary colors. The objects were approximately 12 X 8 X 6 cm and weighed approximately 115 g. Each object was constructed from 2-3 primary shapes (see Figure 1). When objects were acted upon, their shape changed, and each action was unique to each object. For example, one object required 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. Each object made an associated sound when they were acted upon, and it was the sound produced that was of specific interest in this study.

### Training Procedure

The participant and an experimenter sat across from each other at a table. Each had an object set (five objects) randomly lying 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 to engage 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 specific actions. The referee demonstrated the action associated with each of the participant's objects; the experimenter then demonstrated the action associated with each of her objects. Five cards, colored as the monochromatic colors of each object, were placed in a bag. The child was instructed to draw a card, find the toy whose color matched that of the card, and perform the appropriate action on the object 5 times. If required, the referee aided the participant in choosing the correct object and performing the correct action. During learning, the experimenter would say “Look at this, it does this (perform the action)” but did not draw attention to the sound the object made. Actions were not labeled in this study. After drawing all of the cards, the child was rewarded with a sticker. The experimenter drew cards and performed actions on her objects just as the participant had done. During this passive observation condition, the experimenter said “Now I am going to pick this object...Look at what it does!”. The participant and experimenter

alternated turns of the game. Thus, the children participated in both ‘active’ (self generated) and ‘passive’ (observing) conditions. This process was repeated 5 times total, resulting in the experimenter and participant interacting with each of their objects 25 times (5x5 active objects + 5x5 passive objects = 50 total exposures). After completing the game, the referee suggested cleaning up the toys before continuing on to the fMRI session. The participant was instructed to demonstrate the action on each object at least one more time before handing it over to the referee to be put away. The experimenter did the same with the ‘passive’ set of objects. Duration of session was approximately 20-30 minutes.

### **fMRI Test Stimuli**

Auditory stimuli consisted of sounds the 10 learned objects made during the training session along with sounds of 5 unlearned objects. These sounds were pre-recorded and presented to the participants through headphones in the fMRI facility. They were the natural sounds that each action made. Visual stimuli were videos of the actions being performed on learned objects as well as 5 videos of similar, unlearned objects. The videos depicted the objects from a variety of planar (axis of elongation was 0 degrees from the observer) and  $\frac{3}{4}$  (axis of elongation was 45 degrees from the observer) viewpoints. The main shapes of the object could be seen in every video. Auditory and visual stimuli were also combined at times to where a visual video was shown in conjunction with the corresponding auditory sound.

### **Testing Procedure**

After screening and informed consent given by the parent, all participants were acclimated to the MRI environment by watching a cartoon in a 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 (see James, 2010). 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. One to three runs were administered depending on the comfort of the participant. During functional scanning, auditory and visual stimuli were either presented separately or combined - where the appropriate sounds and object videos were linked together. Each run consisted of 6 blocks of stimuli (2 actively learned, 2 passively learned, and 2 novel objects). Blocks were 18-20 seconds long with 10-second intervals between blocks. Each run began with a 10-second rest period and ended with a 20-second rest period. This resulted in runs that were just under 3.5 minutes long. Each presentation of individual stimuli lasted for approximately 2 seconds. During auditory presentation runs, the participant was required to passively listen to the stimuli – no visual stimuli were presented. The participants were told to listen to the sounds that they hear. During the visual perception runs, the participant was required to passively view the videos – no auditory stimuli were presented, participants were instructed to simply watch the videos. During combined auditory and visual runs, the participant viewed the videos while the sound associated with the object played in tandem. Again, participants were instructed to watch the videos-in all three run types, no reference to

the study session was made in the instructions. Neural activation, measured by the BOLD (Blood Oxygen Level Dependant) 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 Acquisition

Imaging was performed using a 3-T Siemens Magnetom Trio whole body MRI system and a phased array twelve 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 x 22 x 9.9 cm, with an in plane resolution of 64 x 64 pixels and 33 slices per volume that were 4mm thick with a 0 mm gap among them. The resulting voxel size was 3.0 mm x 3.0 mm x 4.0 mm. Functional data underwent slice time correction, 3D 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 x 1 mm x 1 mm using trilinear interpolation. High-resolution T1-weighted anatomical volumes were acquired prior to functional imaging using a 3D Turbo-flash acquisition (resolution: 1.25 mm X 0.62 X 0.62, 128 volumes).

### Data analysis procedures

Whole-brain Group contrasts were performed on the resultant data. The functional data were analyzed with a 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. In addition, the movement made by the participants along 3 axes was also included as predictors in the analysis. Any functional data that exceeded 5mm of motion on any axis were excluded prior to analyses. This criterion resulted in excluding 2 blocks of data from one participant and 1 block of data from 3 participants. The attrition rate for this study was 4 individuals who were unable to commence the fMRI portion of the experiment. Exclusion of these data does not significantly alter the power of the present analyses. Data were transformed into a common stereotactic space (eg. Talairach & Tournoux, 1988) for group-based statistical analyses. Direct contrasts of BOLD activation were performed on the group between active sounds and passive sounds that were learned (new sounds were used as a baseline – see below for more detail). In addition, contrasts between activation during perception of objects that were learned actively vs passively were analyzed as well as the combined seeing + hearing the video presentations.

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  $p < .05$ ,



corrected with the False Discovery Rate method (Benjamini, 1995), with a cluster threshold of 270 contiguous 1 mm isometric voxels. (2) Peak statistical probability within a cluster at least  $p < .001$ , corrected.

## RESULTS

### Auditory sound perception

To localize brain regions that were recruited more after active learning than passive learning, a contrast was performed between actively learned sounds and sounds learned through passive observation. Novel sounds were used as a comparison, such that each contrast was structured as follows: (Active sound-novel sound) – (passive learned sound-novel sound). Hearing actively learned sounds resulted in significantly greater activation than passively learned sounds in the middle frontal gyrus bilaterally (Figure 2a), the left precentral gyrus and medial frontal gyrus (Figure 2a), the cingulate gyrus and precuneus (Figure 2b), the anterior middle frontal gyrus (Figure 2b) the insula bilaterally (Figure 2c), the right superior temporal gyrus (Figure 2c) and the bilateral fusiform gyri (Figure 2d).

### Visual object perception

When contrasting actively learned with passively learned objects while participants viewed videos of the objects and actions (active learned video-novel action) – (passive learned action-novel action), three brain regions were recruited significantly more during active than passive learning: The left inferior parietal lobule (Figure 3a), the right superior frontal gyrus (Figure 3a), and the left fusiform (Figure 3b).

### Seeing and hearing learned objects

To determine activation to these conditions, we performed a contrast between (active see +hear-novel see+hear) – (passive see+hear-novel see+hear). In the see + hear condition, only the fusiform gyrus was more active, bilaterally, for the actively learned objects than the passively learned objects (Figure 4). There were no regions where passive observation resulted in greater activation than active exploration during learning.

## DISCUSSION

As organisms that are situated in a given environment and interact with the world through our bodies, it stands to reason that we learn most efficiently when we are able to interact manually with objects. Mounting evidence supports the idea that active interaction leads to better learning in both adults (Harman et al., 1999; James et al., 2002; James, Humphrey & Goodale, 2001; Engelkamp & Zimmer, 1994; Masumoto et al., 2006, Butler & James, 2011) and children (James, 2010, James & Swain, 2011; Longcamp et al., 2005). However, the neural mechanisms that underlie active learning are not well understood. In adults, we know that a history of actions with objects results in activation of the motor system during visual perception – the very system that is used during learning. This ‘re-activation’ is thought to underlie the advantageous affects of active learning. That is, object representations incorporate the motor movements that are used during initial learning, and these motor programs, or plans, are active again during visual perception of the learned objects. Here we

demonstrate four important findings that contribute to this body of knowledge. First, in children, learning about objects actively results in very different brain system recruitment than observing others act. Second, sounds learned actively result in motor, auditory, and *visual* system recruitment more than passive observation upon subsequent auditory presentation. Third, videos of objects that are learned actively result in motor and visual system recruitment more than passively learned objects, and fourth, when given all learned information about objects through multi-sensory presentation, active learning results in greater visual system recruitment than passive observation. We will discuss each of these results in turn.

### Sound perception

Perception of learned sounds is different depending on whether or not those sounds are learned through active engagement with an object or through passively observing another acting on an object. Specifically, the parts of the motor system, including the middle frontal gyrus (premotor areas), the precentral gyrus (primary motor strip), and the cingulate cortex are recruited after active learning. These regions have previously been shown to be recruited during auditory perception after active learning of verbs in children (James & Swain, 2011). As such, these results provide evidence that not only does language recruit motor regions after active learning, rather, *any* association formed between and auditory sound and an action will result in motor system re-activation upon subsequent auditory only presentation. In addition, the insula, an area regarded as involved with sensori-motor processing, was also engaged during auditory presentation after active learning, but not after passive observation. Previous work has shown that the activation of the insula is related to the experience of actions being performed by oneself Farrer & Frith (2002). The activation of this region in the current study suggests reactivation of sensori-motor processes during auditory perception. This mirrors previous work showing that newly learned sound-action associations led to activation of the insula in adults during subsequent auditory perception (Mutschler et al., 2007). This, however, is the first study to show that this region reactivates in children. Being that the right superior temporal gyrus was recruited as a result of active learning suggests that the auditory association regions (STG) were also involved in reactivation more after active learning. Thus, a reactivation of the sensory system involved during encoding occurred- a result that has not yet been shown using the auditory modality.

Interestingly, and most surprisingly, the visual cortex was also recruited during auditory perception after active learning. In other studies, this region of the fusiform has been shown to be engaged after active learning more than passive learning, but only using visual perception during test (James, 2010). However, in the previous research testing auditory verb processing, this region was not active for active learning more than passive observation (James & Swain, 2011), nor was it active during verb perception in children (James & Maouene, 2009). Thus, it seems that sound perception may link to visual object representations more than auditory verb perception. Perhaps the sounds are linked more to the visual information than are verbs, but understanding why this is the case requires further research. We can speculate that perhaps active learning helps to create a stronger association between visual and auditory information. Indeed, we have recently shown that audiovisual associative recognition can be enhanced by active learning (Butler & James, 2011).



## Visual processing

During visual presentation of learned objects, active learning resulted in a recruitment of the IPL more than did passive learning. This is a region associated with grasping in humans and monkeys (eg. Culham & Valyear, 2006) but also and perhaps more relevantly, with tool use (Johnson-Frey, 2004). If our novel objects are represented as known tools after active interaction, then activation in this area is not surprising. Possibly the role of the inferior parietal lobule is not involved in tool representation as much as it is involved in action representation – being located close to parietal regions associated with grasping. Videos associated with actions may recruit regions that are associated with the actual actions, whereas hearing sounds may not activate the actual action patterns associated with interaction, but rather the sensori-motor representation associated with the frontal cortices. The IPL has been implicated previously in action observation that is specific to actions performed by the observer (Clavo-Merino et al., 2005).

In addition, re-activation of the visual cortices was apparent after active learning more than passive observation. This finding supports and extends previous work showing visual system reactivation after active versus passive learning (James, 2010).

Interestingly, visual perception did not recruit auditory cortices, suggesting that perceiving the videos did not access the auditory sounds that the videos produced during training. This is different from the finding that hearing sounds did activate the studied visual images associated with them (see above). This unidirectional re-activation is a novel finding, and one that requires further study.

## Multimodal presentations

The only regions more active after active engagement with objects than passive observation when viewing multimodal presentations was a large section of the visual association cortex, namely, the fusiform gyrus. Again, we see that this region is sensitive to active learning, suggesting that the fusiform receives input from motor regions that are used during learning. We have found this previously with young children (James, 2010) and adults (Butler & James, 2011; James & Atwood, 2009). Sensori-motor processing in the fusiform gyrus has been shown to emerge during perception of some forms that we have motor experience with—such as letters (James & Gauthier, 2006; James et al., 2005). However, the exact role of this region for the integration of sensori-motor signals is not known. One study found that the fusiform was recruited during the integration of visual and haptic information (Kim & James, 2010), a result that may reflect the role of the motor system in haptics and its integration with visual processing. We have hypothesized that motor efferent signals are integrated into visual representations after motor experience, and the end point of these efferents may be the fusiform gyrus (James, 2010; James & Gauthier, 2006).

Therefore, self-generated actions were required for the emergence of motor system recruitment during auditory processing in the developing brain. 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 (eg. Rizzolatti & Craighero, 2004), however, our contrasted conditions both observe actions at test, thus removing any unique affect of this behavior.

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. In other words, covert enactment of the motor patterns associated with the actively learned objects may have recruited the brain regions in this study (i.e. 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 and 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 short coming of fMRI blocked designs (eg. Hauk et al., 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. In addition, we provide the first evidence that shows that sounds associated with objects also recruit these sensori-motor pathways in the young child. Providing such developmental information is important for understanding the cascading effects that certain experiences have on neural response patterns, and therefore, on human cognition.

Furthermore, this work allows us to come closer to understanding the role of the frontal, parietal and occipital systems for object and sound 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. 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.

## ACKNOWLEDGEMENTS

The authors wish to acknowledge all the children and their parents who participated in this study, and also, our MR technicians, Thea Atwood, Rebecca Ward, Ben Pruce, and Chris Chung. Thanks to Linda B. Smith, Susan S. Jones and Thomas W. James for helpful comments on earlier versions of this manuscript. This work was partially

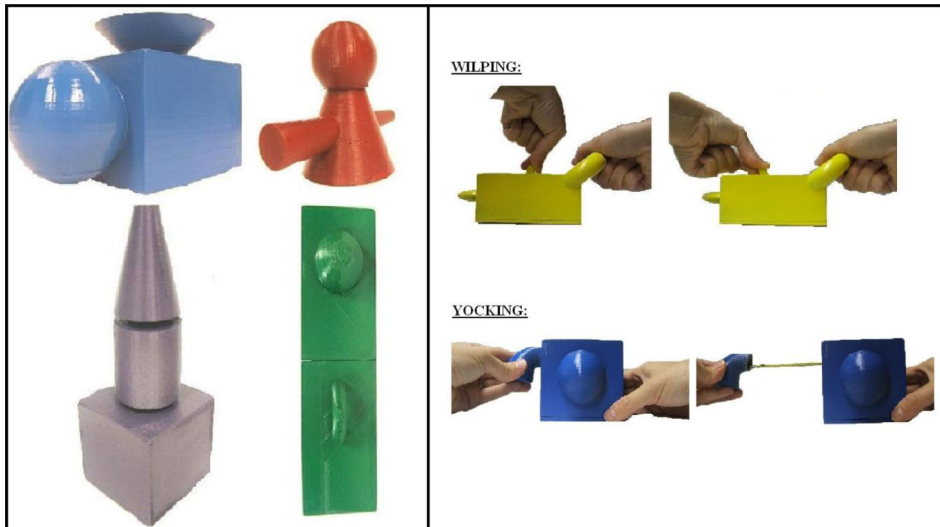
supported by the Lilly Endowment Fund MetaCyte initiative and the Indiana University Faculty Support program to KHJ.

## 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]
- Benjamini, Yoav; Hochberg, Yosef. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)* (Blackwell Publishing). 1995; 57(1):125–13.
- Bertenthal BI, Campos JJ. New directions in the study of early experience. *Child Development*. 1987; 58:560–567. [PubMed: 3608640]
- Bialystok, E. *Language processing in bilingual children*. Cambridge University Press; Cambridge, UK: 1991.
- Bushnell EW, Boudreau JP. Motor development and the mind: The potential role of motor abilities as a determinant of aspects of perceptual development. *Child Development*. 1993; 64:1005–1021. [PubMed: 8404253]
- Butler AJ, James TW, James KH. Enhanced multisensory integration and motor reactivation after active motor learning of audiovisual associations. *Journal of Cognitive Neuroscience*. 2011 Early access, posted on-line March 31, 2011. doi:10.1162/jocn\_a\_00015.
- Calvo-Merino B, Glaser DE, Grezes J, Passingham RE, Haggard P. Action Observation and Acquired Motor Skills: An fMRI Study with Expert Dancers. *Cerebral Cortex*. 2005; 15(8):1243–1249. [PubMed: 15616133]
- Chao LL, Martin A. Representation of man-made objects in the dorsal stream. *Neuroimage*. 2000; 12:478–484. [PubMed: 10988041]
- Chomsky, N.; Ronat, M. Eg. The New Press; NY, NY: 1998.
- Culham JC, Valyear KF. Human parietal cortex in action. *Current Opinions in Neurobiology*. 2006; 16:205–212.
- Dinstein I. Human cortex: Reflections of mirror neurons. *Current Biology*. 2008; 18:R956–R959. [PubMed: 18957251]
- Dinstein I, Hasson U, Rubin N, Heeger DJ. Brain areas selective for both observed and executed movements. *Journal of Neurophysiology*. 2007; 98:1415–1427. [PubMed: 17596409]
- Dinstein I, Gardner JL, Jazayeri M, Heeger DJ. Executed and observed movements have different distributed representations in human AIPS. *Journal of Neuroscience*. 2008; 28:11231–11239. [PubMed: 18971465]
- Dinstein I, Thomas C, Behrmann M, Heeger DJ. A mirror up to nature. *Current Biology*. 2008; 18(1):R13–R18. [PubMed: 18177704]
- Farrer C, Frith CD. Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *NeuroImage*. 2002; 15:596–603. [PubMed: 11848702]
- Gibson, EJ. *Principles of perceptual learning and development*. Appleton-Century-Crofts; East Norwalk, CT: 1969.
- Greces J, Decety J. Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia*. 2002; 40(2):212–222. [PubMed: 11640943]
- Harman KL, Humphrey GK, Goodale MA. Active manual control of object views facilitates visual recognition. *Current Biology*. 1999; 9(22):1315–1318. [PubMed: 10574764]
- Hauk O, Johnsrude I, Pulvermüller F. Somatotopic 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.

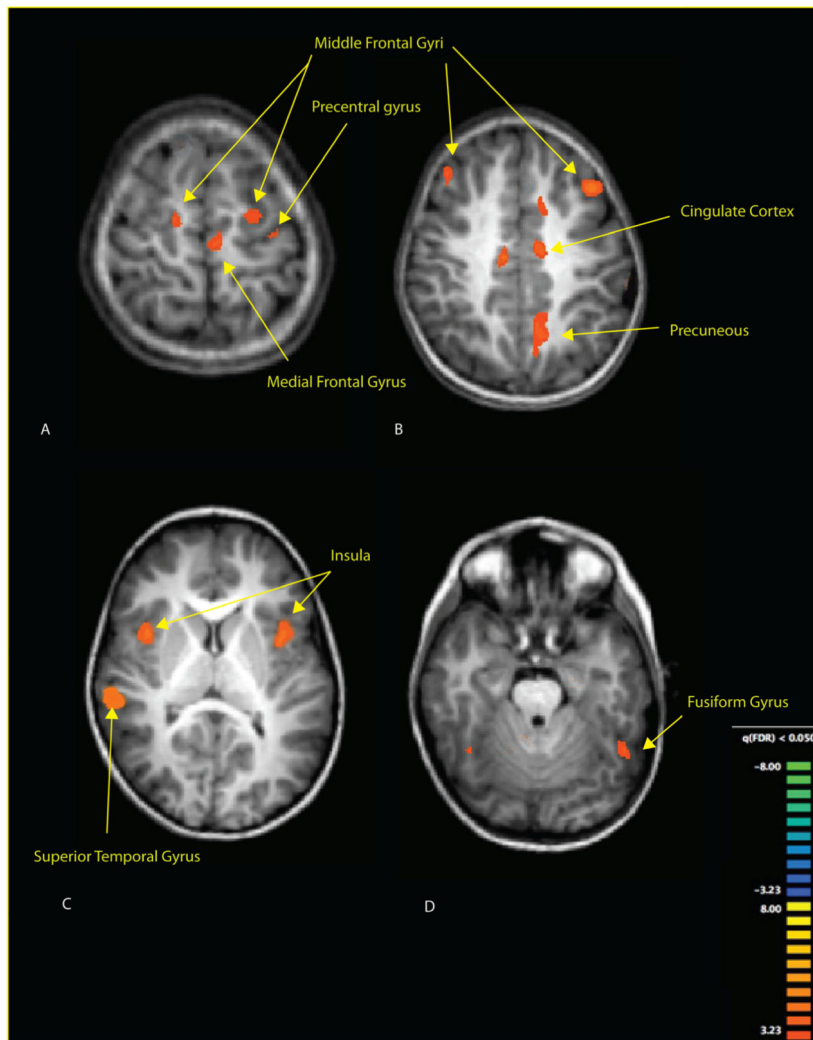
- Huttenlocher J, Smiley P, Charney R. Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review*. 1983; 90:72–93.
- Iacoboni M. Neural mechanisms of imitation. *Current Opinion in Neurobiology*. 2005; 15:632–637. [PubMed: 16271461]
- James KH. Sensori-motor experience leads to changes in visual processing in the developing brain. *Developmental Science*. 2010; 13(2):279–288. [PubMed: 20136924]
- James KH, Gauthier I. Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia*. 2006; 44:2937–2949. [PubMed: 16920164]
- James KH, Humphrey GK, Goodale MA. Manipulating and recognizing virtual objects: Where the action is. *Canadian Journal of Experimental Psychology*. 2001; 55(2):11–120.
- James KH, Humphrey GK, Vilis T, Baddour R, Corrie B, Goodale MA. Learning three-dimensional object structure: A virtual reality study. *Behavioral Research Methods, Instruments, and Computers*. 2002; 34(3):383–390.
- James KH, James TW, Jobard G, Wong C-N, Gauthier I. Letter processing in the visual system: Different activation patterns for single letters and strings. *Cognitive, Affective and Behavioral Neuroscience*. 2005; 5(4):452–466.
- James KH, Maouene J. Auditory verb perception recruits motor systems in the developing brain: An fMRI investigation. *Developmental Science*. 2009:F26–F34. [PubMed: 19840036]
- James KH, Swain SN. Only self-generated actions create sensori-motor systems in the developing brain. *Developmental Science*. 2011; 14:1–6. [PubMed: 21159082]
- Jeanerrod M. Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage*. 2001; 14:103–109.
- Johnson-Frey SH. The neural basis of complex tool use in humans. *Trends in Cognitive Science*. 2004; 8:227–237.
- Kim S, James TW. Enhanced effectiveness in visual-haptic object-selective brain regions with increasing stimulus salience. *Human Brain Mapping*. 2010; 31:678–693. [PubMed: 19830683]
- Longcamp M, Anton J, Roth M, Velay J. Premotor activations in response to visually presented single letters depend on the hand used to write: a study on left-handers. *Neuropsychologia*. 2005; 43(12):1801–1809. [PubMed: 16154456]
- MacWhinney, B. *The Emergence of Language*. Lawrence Erlbaum Associates; 1999.
- Masumoto K, Yamaguchi M, Sutani K, Tsuneto S, Fujita A, Tonoike M. Reactivation of physical motor information in the memory of action events. *Brain Research*. 2006; 1101(1):102–109. [PubMed: 16782071]
- Mutschler I, Schulze-Bonhage A, Glauche V, Demandt E, Speck O, Ball T. A rapid sound-action association effect in human insular cortex. *Public Library of Science One*. 2007; 2:1–9.
- Needham A, Barrett T, Peterman K. A pick-me-up for infants' exploratory skills: Early simulated experiences reaching for objects using 'sticky mittens' enhances young infants' object exploration skills. *Infant Behavior and Development*. 2002; 25:279–295.
- Piaget, J. *The Grasp of Consciousness: Action and concept in the young child*. Routledge and Kegan Pau; London: 1977.
- 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]
- Rizzoltti 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(4):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*. Thieme Medical Publishers; New York: 1988.

- Thelan, E.; Smith, LB. *A Dynamic Systems Approach to the Development of Cognition and Action*. MIT Press; Cambridge, MA: 1994.
- Weisberg J, van Turenout M, Martin A. A neural system for learning about object function. *Cerebral Cortex*. 2007; 17(3):513–521. [PubMed: 16581980]

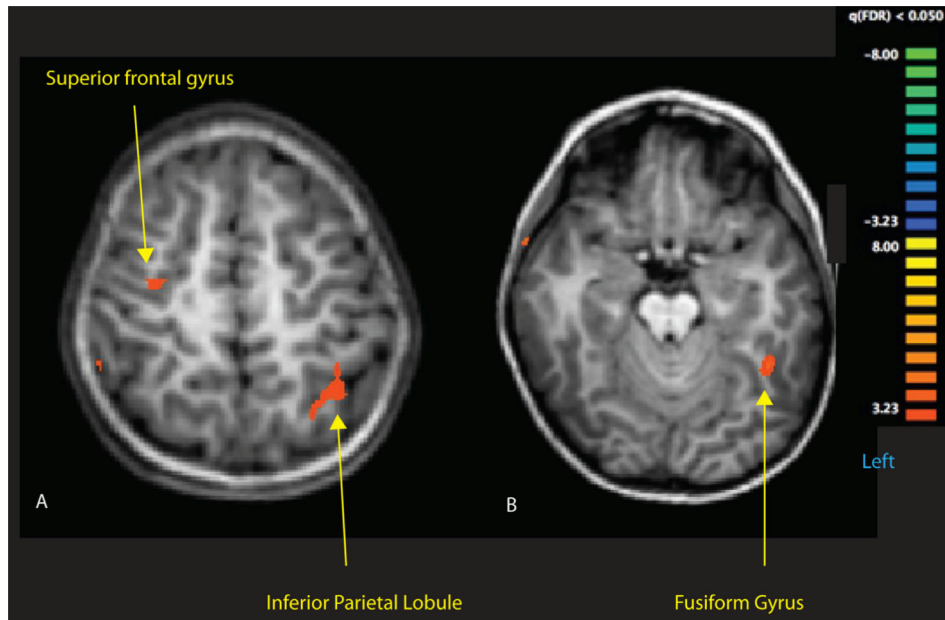


**Figure 1.** Examples of 3-D plastic objects used during training and testing. Images are taken from video clips from the MRI testing stimuli.



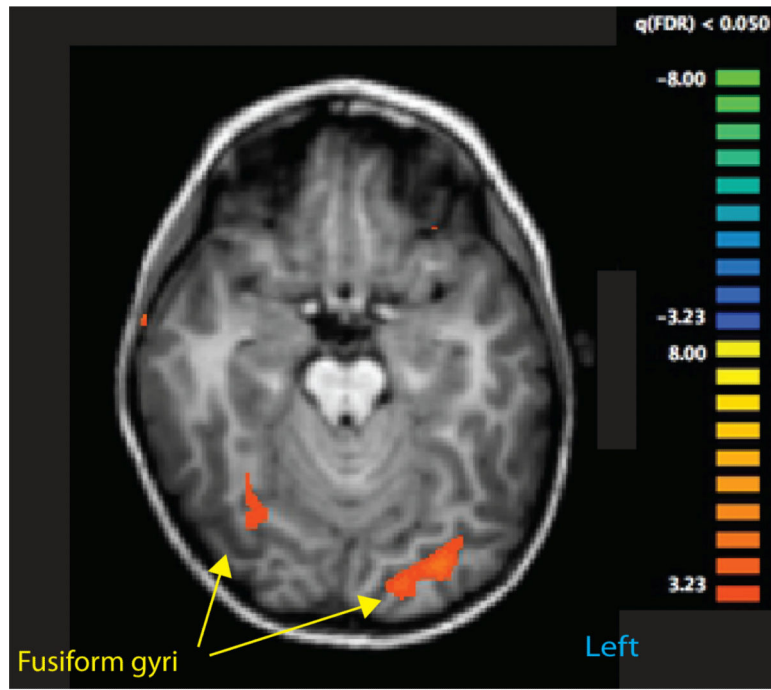


**Figure 2.** Recruitment of brain regions during sound perception that are significantly more active after learning sounds with active interaction than by passive experience. Left hemisphere is on the right. Coordinates reported for all regions are according to the conventions of Talairach & Tourneaux, 1988). 2A) Middle frontal gyri (-23,-9,60; 19,-9,60) Precentral gyrus (-30,-18,60) and Medial frontal gyrus (-5,-21,60). 2B) Middle frontal gyri-anterior (-43,22,40; 39,29,40), Cingulate cortex (-7,22,43; 5, 22,43) and (-13,-10,42), and the Precuneus (-11,-52,42). 2C) Activation in the Insula (35,6,6; -39,12,6) and the Superior Temporal Gyrus (52,-23,8). 2D) Activation in the Fusiform gyrus (-49,-51,-21). No regions were significantly more active after passive learning.



**Figure 3.**

Regions that are significantly more active when viewing videos of objects and actions after active exploration during learning vs. passive observation during learning. 3A) Superior Frontal gyrus (29,-11,51) and the Inferior parietal lobule (-37,-50,51). 3B) Fusiform gyrus (-44,-43,-12). No regions were more active after passive observation in this contrast.



**Figure 4.** Significantly greater activation in the see and hear video conditions after active exploration vs. passive observation during learning. The only regions more active here are the fusiform gyri ( $-19, -88, -17$  to  $-32, -80, -17$ ) and  $29, -61, -17$ ).

**Table 1**

Object label, action description and sound description of stimuli used in the training sessions.

<b>Object Label (not given to subject)</b>	<b>Action description</b>	<b>Sound description</b>
Blue Yocking	Pulling out part	Paper cutter
Green ratching	Putting part onto object	Clinking
Purple sprocking	Pushing down part	Squeeking
Red tilfing	Pulling apart	Scraping plastic
Yellow wilping	Pulling up part	Squishing
Blue quaning	Rotating part	Stirring ice
Red luing	Pushing button	Glass breaking
Purple manuing	Pulling apart	Radio white noise
Red nooping	Pushing two parts together	Metal scraping
Yellow panking	Pushing part into hole	Toy rattle
Purple scoching	Closing/opening	Buzzing
Blue tooping	Pushing in part	Swirling water
Green laking	Pushing down top	Scissors cutting
Red patchoing	Shaking object	Bong bong sound
Yellow sapping	Bouncing object	Rattle snake
Purple lewing	Putting two parts into holes	Clicking
Blue dewping	Pushing object up	Screeching
Red toppzing	Stirring stick on object	Tweeting
Purple hanning	Rotating object	Tires screeching
Yellow silking	Pulling up parts	Twigs breaking
Green halking	Pushing down part	Barking
Blue razzing	Closing two sides together	Chairs scraping on floor
Red bulking	Pushing one then another button	Whistling
Purple gupping	Twirling piece on object	Siren