

# Crossmodal enhancement in the LOC for visuohaptic object recognition over development



R. Joanne Jao <sup>a,b,\*</sup>, Thomas W. James <sup>a,b,c</sup>, Karin Harman James <sup>a,b,c</sup>

<sup>a</sup> Cognitive Science Program, Indiana University, Bloomington, USA

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

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

## ARTICLE INFO

### Article history:

Received 9 December 2014

Received in revised form

5 August 2015

Accepted 7 August 2015

Available online 10 August 2015

### Keywords:

Multisensory

Object recognition

Crossmodal

Vision

Haptics

fMRI

Development

## ABSTRACT

Research has provided strong evidence of multisensory convergence of visual and haptic information within the visual cortex. These studies implement crossmodal matching paradigms to examine how systems use information from different sensory modalities for object recognition. Developmentally, behavioral evidence of visuohaptic crossmodal processing has suggested that communication within sensory systems develops earlier than across systems; nonetheless, it is unknown how the neural mechanisms driving these behavioral effects develop. To address this gap in knowledge, BOLD functional Magnetic Resonance Imaging (fMRI) was measured during delayed match-to-sample tasks that examined intramodal (visual-to-visual, haptic-to-haptic) and crossmodal (visual-to-haptic, haptic-to-visual) novel object recognition in children aged 7–8.5 years and adults. Tasks were further divided into sample encoding and test matching phases to dissociate the relative contributions of each. Results of crossmodal and intramodal object recognition revealed the network of known visuohaptic multisensory substrates, including the lateral occipital complex (LOC) and the intraparietal sulcus (IPS). Critically, both adults and children showed crossmodal enhancement within the LOC, suggesting a sensitivity to changes in sensory modality during recognition. These groups showed similar regions of activation, although children generally exhibited more widespread activity during sample encoding and weaker BOLD signal change during test matching than adults. Results further provided evidence of a bilateral region in the occipitotemporal cortex that was haptic-preferring in both age groups. This region abutted the bimodal LOTv, and was consistent with a medial to lateral organization that transitioned from a visual to haptic bias within the LOC. These findings converge with existing evidence of visuohaptic processing in the LOC in adults, and extend our knowledge of crossmodal processing in adults and children.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Vision is the dominant perceptual modality in humans. This is especially true for object recognition – a ubiquitous and highly important cognitive function – and particularly for recognizing objects based on their three-dimensional shape. However, it has been demonstrated that, in the absence of vision, accurate recognition of objects based on shape cues can be accomplished using haptic input alone (Klatzky et al., 1987; Lederman and Klatzky, 1987, 1990, 1993; Norman et al., 2004, 2008). In addition to these unisensory findings, there is evidence that visual and haptic inputs are effectively combined to enhance recognition performance (Kim and James, 2010; Kim et al., 2012), and that the

haptic modality can successfully prime the visual modality and vice versa (Easton et al., 1997; James et al., 2002). Finally, it has been shown that objects can be successfully matched when studied either visually or haptically, and then tested with the other modality (i.e., crossmodal matching; James et al., 2007; Kassuba et al., 2013; Newell et al., 2001). Combined, this evidence suggests that the visual and haptic systems are tuned to the shape properties of objects, and furthermore, that they share shape information and perhaps even common representations for the purposes of object recognition (Amedi et al., 2005; James, James, Humphrey, et al., 2006; Lacey and Sathian, 2011).

For the past two decades, a body of evidence has been accrued suggesting that regions of the putative visual and haptic cortices are involved in combining visual and haptic information about object shape. This research has increasingly concentrated on multisensory object recognition, and more specifically, on visuo-haptic integration (Amedi et al., 2001, 2002; James et al., 2002; James and Kim, 2010; Kassuba et al., 2013; Lacey and Sathian,

\* Correspondence to: Indiana University, 1101 E. Tenth St., Bloomington, IN 47405, USA.

E-mail address: [rjjao@indiana.edu](mailto:rjjao@indiana.edu) (R.J. Jao).

2011; Lacey et al., 2009; Gentile et al., 2011; Stilla and Sathian, 2008). Several neuroimaging studies have determined that this convergence of visuohaptic information in adults occurs most commonly at two particular brain regions: the lateral occipital complex (LOC; Amedi et al., 2001, 2002; James et al., 2002; Stoesz et al., 2003; Reed et al., 2004; Prather et al., 2004; Pietrini et al., 2004), and the intraparietal sulcus (IPS; James and Kim, 2010; Bodegard et al., 2001; Binkofski et al., 1999; Culham and Kanwisher, 2001; Grefkes et al., 2002; Peltier et al., 2007; Roland et al., 1998; Zhang et al., 2004; Stilla and Sathian, 2008). The LOC is a large collection of regions, including the anterior aspects of the inferior and middle occipital gyri, the posterior aspects of the middle and inferior temporal gyri, and a large section of the occipital and temporal fusiform gyrus. The LOC has been found to respond more to (images of, and three-dimensional forms of) visual presentations of intact objects such as tools, animals, toys, etc., as compared to scrambled versions of the same objects or textures (Amedi et al., 2001; Grill-Spector et al., 2001, 2008; James and Kim, 2010; Kassuba et al., 2011; Kourtzi and Kanwisher, 2001; Malach et al., 1995; Tootell et al., 1996). In addition, the lateral occipital tactile-visual region (LOtv) contained within the complex, activates during haptic presentations of those objects compared to textures (Amedi et al., 2001, 2002; James et al., 2002; Stilla and Sathian, 2008).

The IPS is a one of the primary landmarks of the parietal lobe and runs from the occipitoparietal junction to cross the post-central sulcus. The anterior aspect of the IPS (aIPS) lies near the post-central gyrus and has been commonly reported to activate during haptic shape perception (Bodegard et al., 2001; Culham and Kanwisher, 2001; Peltier et al., 2007; Roland et al., 1998; Stilla and Sathian, 2008; Zhang et al., 2004). Furthermore, the anterior and posterior aspects of the IPS have been shown to activate during the crossmodal matching of two- and three-dimensional objects (Grefkes et al., 2002; Saito et al., 2003), demonstrating a preferential response for overall object shape regardless of sensory modality. The object-responsive IPS has been found to activate not only to common, familiar objects (Amedi et al., 2001, 2002, 2005; Deibert et al., 1999; Reed et al., 2004), but to simple, geometrical shapes as well (Bodegard et al., 2001; Roland et al., 1998). Similarly, the LOC has been found to respond to both familiar and novel shape information (Lacey et al., 2010, 2014). Taken together, this accrual of evidence implicates the LOC and IPS as sites of convergence in which visuohaptic information is processed for the analysis of object shape.

Relative to adult research, there are far fewer neuroimaging studies that have investigated the development of crossmodal perception of visual and haptic information. Yet, a full understanding of the mechanisms of visuohaptic convergence requires an understanding of their development. In fact, understanding how convergence of sensory systems develops may provide unexpected insights into the mechanisms of adult multisensory processing. Behavioral studies that have researched this question suggest that communication within sensory systems generally develops earlier than communication across sensory systems (Bushnell and Baxt, 1999). Additionally, research has shown that the manner in which infants and children haptically explore objects influences concurrent and later visual perception (Ruff, 1984, 1986, 1989; Bushnell and Boudreau, 1993), and more adult-like patterns of visuomotor exploration at 24 months result in an increase in some measures of visual object recognition (James et al., 2014). By 4 to 5 years of age, children's patterns of haptic object exploration begin to appear stereotypically adult-like (Kalagher and Jones, 2011a, 2011b), and intramodal haptic object recognition is highly accurate (Bushnell and Baxt, 1999). Nevertheless, crossmodal visuohaptic recognition abilities at this age are not yet adult-like. At 5 years, children demonstrate poorer crossmodal

performance compared to intramodal visual-to-visual or haptic-to-haptic recognition for novel objects (Bushnell and Baxt, 1999). It is not until 8–10 years that the integration of visual and haptic shape information becomes statistically optimal according to psychophysical discrimination tasks, which suggests a developmental transition prior to this age range (Gori et al., 2008). As such, the behavioral delay in crossmodal processing is thought to be due to the lack of efficiency in integrating or transferring information from one modality to another before 8 years of age.

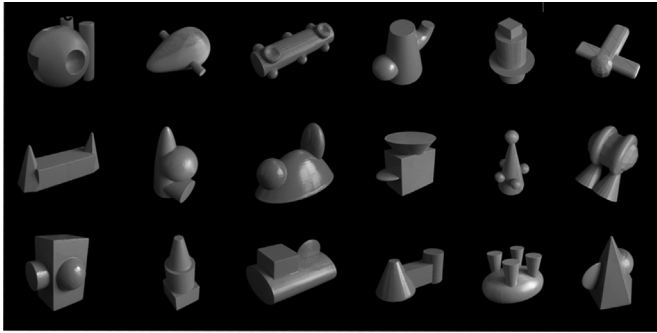
More is known about the neural mechanisms that underlie the developmental progression of visual than of haptic object recognition. For vision, there is evidence that the LOC is generally recruited during visual object perception by 7 years (Grill-Spector et al., 2008; Scherf et al., 2007). This recruitment may be experience-dependent as evidenced by differential LOC recruitment among similarly aged children based on the level of experience with select object classes (James and James, 2013). Thus, although the LOC is recruited for visual object recognition early on in development (i.e., by 7 years of age), the specific pattern of recruitment of the LOC and of the surrounding cortex continues to develop after 7 years (Grill-Spector et al., 2008; Scherf et al., 2007). In a previous study by our lab, we found that unisensory haptic object preference was adult-like by 4 to 5.5 years, whereas visual object preference continued to increase into young adulthood at which point it was quite visually dominant (Jao et al., 2014). A key conclusion of these findings was that the apparent level of maturation of function within a particular neural substrate is task-dependent.

As such, the goal here was to examine visuohaptic interactions with respect to crossmodal matching. We implemented a delayed match-to-sample paradigm and measured blood oxygen level dependent (BOLD) responses in 7–8.5-year-old children and young adults during crossmodal visuohaptic and intramodal visual and haptic processing of novel objects. Novel objects were used, based on previous findings of behavioral differences during crossmodal and intramodal tasks for unfamiliar objects (Bushnell and Baxt, 1999). The age groups were chosen to encompass the estimated age interval of the developmental transition for visuohaptic integration as indicated by previous behavioral and neural reports (Gori et al., 2008; Grill-Spector et al., 2008; Jao et al., 2014; Scherf et al., 2007). Given this previous work, we predicted a stronger response for crossmodal than intramodal matching in adults, and possibly in children (i.e., the crossmodal matching effect in IPS; see Grefkes et al., 2002). Moreover, we predicted that the marker used to indicate efficient sharing of crossmodal sensory information, namely crossmodal enhancement—as evidenced by increased responses to crossmodal versus intramodal sequentially-matched stimuli, and distinct from multisensory enhancement or multisensory gain (Kim and James, 2010; Kim et al., 2012; also see Stein and Stanford (2008))—would be less pronounced in children, reflecting their decreased ability to integrate visual and haptic information about object shape.

## 2. Methods

### 2.1. Participants

Participants were recruited from two age groups: 7–8.5-year-old children ( $N=11$ , 5 female, mean age=7.8 years,  $\sigma=0.4$  years) and young adults ( $N=10$ , 5 female, mean age=25.8 years,  $\sigma=5.7$  years). Three additional children were tested, but were excluded from analyses due to excessive motion (2) and high variability (1). Participants had normal or corrected to normal vision, had no known history of psychological disorders, and were predominantly right-handed (7–8.5 years: 1 ambidextrous, 1 slight



**Fig. 1.** Subset of novel stimuli used in the present study. Objects are not to scale.

left-preferring, 9 right-preferring; Adults: 10 right-preferring) as measured by the Edinburgh Handedness Inventory (Oldfield, 1971; for a discussion of handedness, see subsection 4.4 *Handedness and the LOC*). All met the criteria for MRI scanning. Written informed consent was obtained from the parents and adult participants, and written informed assent was obtained from the children. Parents were compensated with a gift certificate, children were compensated with a small toy, and adult participants were compensated with \$25. This research was approved by the Indiana University Protection of Human Participants Board.

## 2.2. Stimuli

The stimuli consisted of 20 objects that were explored visually and haptically. Stimuli were novel, three-dimensional, and had rigid bodies that were controlled for texture (i.e., all objects were printed in plastic using a 3-D printer) (Fig. 1). The visual stimuli consisted of gray scale photographs of the haptic stimuli at a typical three-quarters viewing angle against a solid black background to facilitate recognition during visual exploration. Stimuli were further controlled for size to ensure that the younger group of participants could fit both of their hands around them during haptic exploration; objects were no greater than 9 cm along the longest dimension, and no smaller than 2.5 cm along the shortest dimension. Participants did not see or feel the objects prior to the training session that occurred before the imaging session.

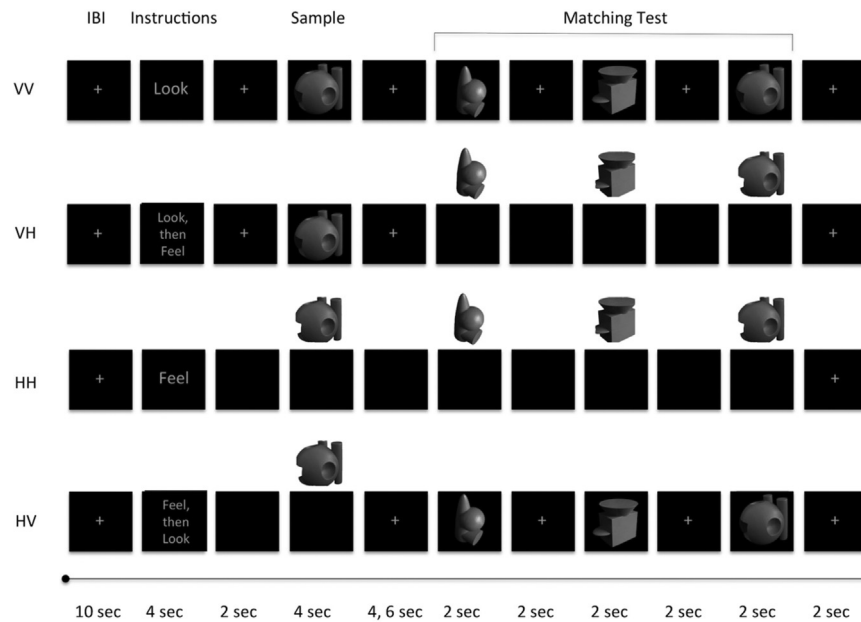
## 2.3. Procedure

After screening and obtaining informed consent from the adults and assents from the children, all participants were acclimated to an MRI environment. Children watched as a short cartoon was played on a screen in the MRI simulator, an artificial MRI environment with the same dimensions and sounds as the actual MRI environment. Participants were then trained in the experiment. They were instructed to lie very still in a supine position, and a lap desk was placed over their midsection. A cape was placed over their torso and arms, and was tucked under their chin. This cape covered the lap desk and allowed the participants to feel the stimuli with their hands without being able to see them. Participants were trained to perform four types of delayed match to sample tasks following the instructions: 1) “Look;” 2) “Feel;” 3) “Look, then feel;” and 4) “Feel, then look.” The delayed match to sample tasks were further separated into two phases: 1) the sample encoding phase; and 2) the test matching recognition phase. The first two types of instructions yielded intramodal matching (i.e., visual-to-visual or “VV,” haptic-to-haptic or “HH”) in which the same modality was used to encode and recognize the stimuli. The latter two types of instructions resulted in crossmodal matching (i.e., visual-to-haptic or “VH,” haptic-to-visual or “HV”) in which one modality was used to encode the stimulus and

another modality for testing recognition as well as crossmodal sharing of object shape information. During the encoding phase for each type of delayed match to sample task, participants were instructed to look at or feel the sample stimulus. They were then tested on intramodal and crossmodal recognition with three objects sequentially, each of which they had to decide if it was the same as or different from the sample stimulus. Participants verbally indicated their responses during the training tasks, and were able to do so correctly with ceiling levels of performance (i.e., participants in each group for each task responded correctly on all trials).

Once the participants were comfortable in this setting and could perform the tasks efficiently, they were introduced to the actual MRI environment. It was decided based on several reasons that behavioral responses would not be recorded during the MRI scan. First, both children and adults show a reliable preference for two-handed haptic exploration of 3D objects, particularly with regard to shape recognition (Lederman and Klatzky, 1987; Kalagher and Jones, 2011a, 2011b). To include a button press response would restrict exploration to the unnatural single-handed mode, which would likely be more distracting for children than adults. Alternatively, exploration could be two-handed, followed by a button press, but finding the location of the button without being able to locate it visually would be difficult, as would returning to the start position for haptic exploration. This alternative would likely be more challenging for children than adults. Outside the MRI environment, verbal responses would be an excellent candidate to combine with two-handed exploration, but collecting verbal responses within the MRI scanner introduces considerable artifacts into the BOLD signal. Another possibility would be foot-button presses as responses; however, this type of response produces significantly more head motion in adult participants (Kim and James, 2010; unpublished data) and was dismissed as an option because of the likelihood of extreme head movements it may have produced in children. These considerations were in addition to the fact that the delayed match-to-sample task is already relatively cognitively challenging, especially when considering the working memory load involved in keeping one of four instructional contexts in mind while performing that task. The final choice to eliminate behavioral responses from the MR experimental protocol was based on our previous developmental neuroimaging findings showing that forcing children to perform tasks that are too cognitively or attentionally demanding results in early withdrawal or, when they do not withdraw, in increased head and body motion. Although this approach may not be completely ideal and may potentially be criticized for being unable to ensure the engagement of neural processes without recorded behavioral responses from the MRI scan, it was nevertheless the most reasonable means available to examine the development of crossmodal processing using delayed match-to-sample tasks in this difficult-to-image population. Based on the rationale described above, performance was therefore measured beforehand during training instead of during the MR protocol, and indicated that children could perform as well as adults. In terms of ensuring an equal level of task compliance between children and adults, an experimenter was present in the MR room during the scan to monitor the participant's behavior, including compliance and body and head movement. Additionally, instructions were provided before each functional run, and participants verbally confirmed before each run that they understood what was expected as well as after each run that they were indeed performing the task.

Once in the MRI, participants were again given the instructions, and the lap desk and cape were placed over their midsection. All visual stimuli were back-displayed via a Mitsubishi XL30 projector onto a screen located behind the participants in the bore of the MRI; this screen was viewed through a mirror that was placed on



**Fig. 2.** Graphical depiction of the fMRI mixed event-related/block design. This figure demonstrates the timing protocol of a single trial set for each condition.

top of the head coil. Visual stimuli were presented using SuperLab Pro 2.0.4 software from an Apple MacBook laptop. Haptic stimuli were exchanged by the experimenter in the MRI, and were attached to the lap desk with Velcro so that participants could not lift the objects off of the surface during exploration. Participants viewed the visual stimuli with both eyes, and explored the haptic stimuli using both hands.

A high-resolution anatomical scan was first acquired, which for children, occurred as they watched a cartoon. Upon completion of this scan, four functional scans were acquired. Participants were tested using a mixed event-related/block design that involved intramodal visual encoding and recognition of 2D images of the stimuli, intramodal haptic encoding and recognition of the 3D stimuli, as well as crossmodal visual-to-haptic and haptic-to-visual encoding and recognition of the stimuli. Examples of all 4 tasks are depicted in Fig. 2. For each type of task, instructions were presented for 4 s, followed by a 2 s inter-stimulus-interval (ISI). The sample stimulus was then presented for 4 s. After a variable delay of 4 or 6 s in between sample and test, the test stimuli were presented sequentially for 2 s with a 2 s ISI between each. An inter-trial-interval (ITI) of 2 s, during which participants viewed a gray fixation cross, separated trial sets—a set comprised the instructions, sample presentation, delay, and test presentations. Finally, an inter-block-interval (IBI) of 10 s was presented at the beginning and end of each run. It is important to note that in order to dissociate the relative contributions of sample and test during intramodal and crossmodal matching, trial sets were not evaluated as a compound event (for a brief critique, see Kassuba et al. (2013)). Rather, the variable delay allowed for later deconvolution of the neural signal into the two phases.

Throughout the functional scanning session, participants viewed a black background on which the visual instructions, visual stimuli, and gray fixation cross (ISI, ITI, IBI) were presented, or a blank screen was displayed. Since instructions were given visually, participants were instructed to keep their eyes open during haptic exploration; therefore, the screen remained blank during and between presentations of haptic stimuli.

The experiment consisted of 32 event-related sample trials (8 per task) and 96 blocked test trials (24 per task) in total, separated into 4 runs of approximately 4-minute-long functional scanning (230 s, 115 volumes) with 8 trial sets per run. Trials were pseudo-

randomized such that each trial set consisted of at least one match and one mismatch between the test and sample stimuli. Each run comprised of either VV and VH conditions (visual encoding tasks) or HH and HV conditions (haptic encoding tasks). This was to ensure that the modality used during the encoding phase was consistent across the run, and to minimize task-switching errors, particularly by the children. Tasks were counterbalanced across runs, and run order was randomized for each participant. Imaging sessions lasted approximately 30 min. After the scanning was completed, participants were removed from the MRI environment and compensated for their time.

#### 2.4. MRI data acquisition

Imaging was performed using a 3-Tesla Siemens Magnetom Trio whole body MRI system located within the Imaging Research Facility at the Indiana University Psychological and Brain Sciences department. A phased array 12 channel head coil was used to obtain whole-brain functional volumes; these were acquired using a gradient echo planar imaging (EPI) sequence (TE=30 ms, TR=2000 ms, flip angle=70°) for BOLD-based imaging. The field of view was 192 cm with an in-plane resolution of 64 × 64 pixels and 33 slices per volume (3.8 mm thick with a 0 mm gap), which resulted in a voxel size of 3 × 3 × 3.8 mm<sup>3</sup>. Using analysis tools in the BrainVoyager QX™ 2.4 software package (Brain Innovation, Maastricht, Netherlands), functional data underwent slice scan-time correction, 3D motion correction, linear trend removal, and Gaussian spatial blurring (FWHM 6 mm). High-resolution T1-weighted anatomical volumes (resolution: 1.5 × 1.5 × 1.5 mm<sup>3</sup>, 120 sagittal slices) were acquired prior to the functional imaging using a 3-D Turbo-flash inversion recovery sequence. Individual functional volumes were co-registered to the anatomical volumes with an intensity-matching, rigid-body transformation algorithm. Anatomical and functional volumes were normalized to a standard space using an affine transformation based on the 8 parameters of the Talairach reference (Talairach and Tournoux, 1988). For a discussion of the concerns regarding comparisons between children and adult brains normalized to a standard, stereotactic atlas such as the Talairach space, see Appendix B of Wakefield et al. (2013). During normalization, voxels of the functional volumes were re-sampled to 3 mm<sup>3</sup>.

## 2.5. Data analysis procedures

### 2.5.1. Group contrasts

Whole-brain statistical parametric maps (SPMs) were calculated using the BrainVoyager QX™ 2.4 analysis package. Data were transformed into a common stereotactic space (e.g., [Talairach and Tournoux, 1988](#)) for group-based statistical analyses. Functional data were analyzed using a random effects general linear model (GLM) with group as a between-subjects factor. Two predictors were entered in the design matrix for each of the four crossmodal or intramodal conditions. The onset time of the first predictor was based on the sample stimulus presentation time (sample encoding phase) and the onset time of the second predictor was based on the onset of the block of test stimuli (test matching phase). These predictors were convolved with a two-gamma hemodynamic response function. Motion parameters were also included in the design matrix as predictors of no interest. Functional runs with motion estimates exceeding 5 mm on any axis were excluded from the analyses. Although this is a more liberal threshold than is often used in studies with only adults, it was adopted as the criterion here because a stricter criterion would have eliminated most of the child participants (see [Section 2.5.3](#)). This criterion resulted in a total of 42 runs (on average, 3.8 runs per participant) for the children, and 40 runs (4 runs per participant) for the adults.

The whole-brain contrasts were thresholded using a minimum voxel-wise  $p$ -value of  $<.001$  per map, and corrected for multiple tests using a cluster threshold of at least 29 voxels as determined by Monte Carlo simulation using the BrainVoyager QX™ Cluster-size Estimation Plug-in. This plug-in estimates the cluster-size threshold required to produce an  $\alpha <.05$  based on a specific voxel-wise  $p$ -value.

### 2.5.2. Post-hoc region-of-interest selection

A post-hoc Region-of-Interest (ROI) analysis was performed on a group-defined ROI in the LOC to examine the effects of crossmodal and intramodal processing during the test matching phase. To localize this ROI, four whole-brain SPMs were calculated during the sample encoding phase using random-effects GLMs and balanced contrasts. These contrasts compared: a) adult vision to rest ( $VV+VH > 2 \times \text{rest}$ ); b) adults haptics to rest ( $HH+HV > 2 \times \text{rest}$ ); c) 7–8.5 year old vision to rest ( $VV+VH > 2 \times \text{rest}$ ); and d) 7–8.5 year old haptics to rest ( $HH+HV > 2 \times \text{rest}$ ). The LOC ROI was selected as the overlap between all four contrasts in the left hemisphere. Previous studies have demonstrated robust effects in the left hemisphere ([Kim and James, 2010](#); [Kim et al., 2012](#)), and while our results showed no difference between hemispheres, we primarily show the data from the left hemisphere for convenience. It is important to note that the current study did not include a within-modality control condition (e.g., textures; see [Amedi et al. \(2001, 2002\)](#) and [Kassuba et al. \(2011\)](#)). In the absence of such a condition, there are limitations to the specificity of the conclusions that can be made regarding object recognition. That is, without being able to contrast objects with textures within each modality directly, there is a possibility that any effects in the LOC could be related to other processes and not to object recognition. However, the LOC ROI was comparable in terms of anatomical overlap to the LOTv area indicated by a previous study that compared visuohaptic object-selectivity directly (i.e.,  $(VO > VT) \cap (HO > HT)$ ; [Jao et al., 2014](#)). Thus, given the location of this ROI, it appears to reflect a similar functional region of cortex (i.e., LOTv) as indicated by other studies to be involved in visuohaptic object recognition (e.g., [Amedi et al., 2001, 2002](#); [Amedi et al., 2010](#); [Kim et al., 2012](#)), and provides an adequate proxy for LOC ROIs defined using a more standardized task.

A secondary post-hoc ROI analysis was conducted on group-defined ROIs located medially and laterally to the left LOC ROI.

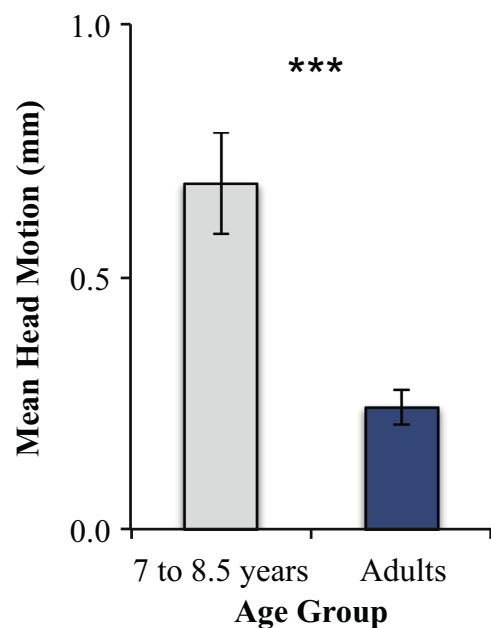
These regions were localized using overlapping difference maps comparing vision to haptics ( $VV+VH > HH+HV$ ) in each group during the sample encoding phase. The medial (fusiform gyrus) and lateral (middle temporal/occipital gyrus) ROIs were examined for differential effects of modality, as well as of crossmodal versus intramodal processing, in relation to the overlap (LOC) ROI during the test matching phase.

BOLD time courses were extracted from these ROIs during the test matching phase for each condition using event-related averaging. BOLD activation values were calculated for each participant for each condition as the mean BOLD signal change that was time locked to the onset time of the block of test stimuli (i.e., between 14 and 16 s post-trial onset, which began with the instructions) to measure activation during the test matching phase. Two dependent measures were of interest, including: a) the test trial type—this measure was based on crossmodal (VH and HV) versus intramodal (VV and HH) matching; and b) the test modality—this measure was based on the sensory modality used to match the test stimulus to the sample stimulus (vision: VV and HV; haptics: HH and VH). For the LOC ROI, BOLD activation values were used as the dependent measure in a  $2 \times 2 \times 2$  split-plot repeated measures analysis of variance (ANOVA) performed in SPSS with group as a between-subjects factor and modality and trial type as within-subjects factors. For the medial, overlap, and lateral ROIs, BOLD activation values were used as dependent measures in a  $2 \times 2 \times 2 \times 3$  split-plot repeated measures ANOVA with group as a between-subjects factor and modality, trial type, and ROI as within-subjects factors.

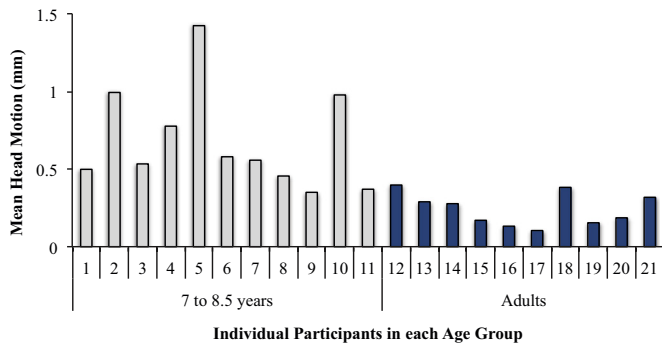
### 2.5.3. Motion tolerance threshold analyses

Motion tolerance threshold analyses were conducted to examine age-related differences in motion, and the impact of motion artifacts on the BOLD signal (see supplementary materials in [Jao et al. \(2014\)](#) for a detailed description).

Results showed that children aged 7–8.5 years did produce greater head motion than adults ( $t(19)=4.280$ ,  $p <.001$ ). The mean motion (mm) for each age group is presented in [Fig. 3](#); the mean head displacement (mm) of individual subjects is presented in [Fig. 4](#). Although adult head motion, measured using mean



**Fig. 3.** Mean head motion (in mm) for each age group. On this and subsequent figures, error bars represent standard error of the mean; significant differences at  $p <.001$  are depicted with \*\*\*,  $p <.01$  with \*\*, and  $p <.05$  with \*.



**Fig. 4.** Mean head motion (in mm) for each individual participant within each age group.

**Table 1**

Number of subjects retained at different motion tolerance thresholds.

Motion threshold (mm)	Total no.	7–8.5 years	Adults
5	21	11	10
4	21	11	10
3	17	7	10
2	17	7	10
1	13	3	10

motion (mm) or maximum head displacement (mm), was predominantly below 1 mm, the number of retainable subjects in the group of 7–8.5 year olds would have decreased substantially as the threshold became stricter (Table 1). Thus, the more liberal

threshold of 5 mm is more inclusive for difficult-to-image child populations, while a stricter criterion would be impractical.

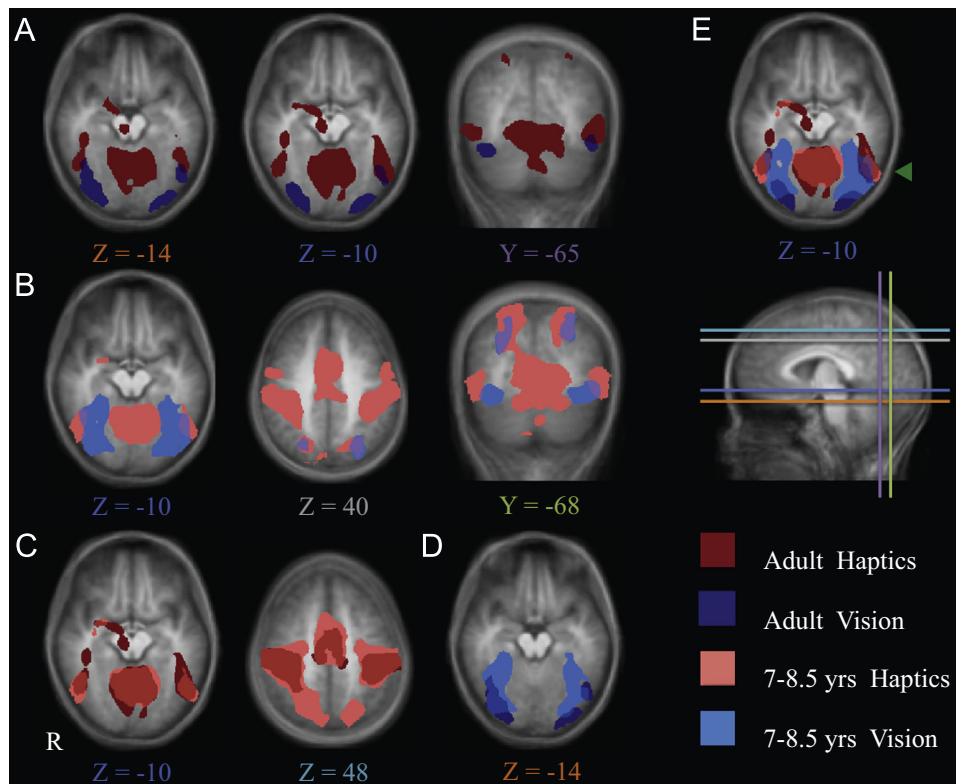
Yet, it must be ruled out that the tendency for children to move their heads more so than adults may have driven the differences in BOLD activation. To address this concern, the BOLD signal change for crossmodal processing—the combined dependent measure of VH and HV that produced the largest effect—was compared to mean motion in each group. Results indicated that the correlation between these two measures was not significant in either age group (7–8.5 year olds: ( $r^2(9)=0.070$ ,  $p=n.s.$ ); adults: ( $r^2(8)=0.172$ ,  $p=n.s.$ )). Based on these findings, it is unlikely that the differences in neural activity can be accounted for by head motion.

### 3. Results

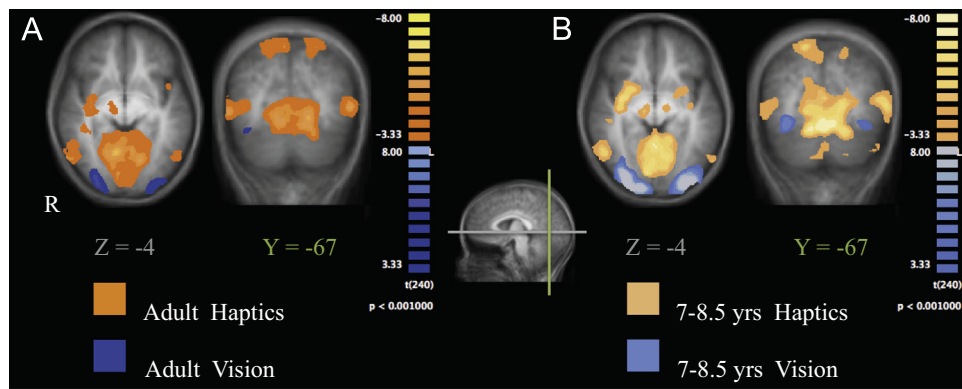
#### 3.1. Sample encoding phase

##### 3.1.1. Overlapping maps of activation

Statistical Parametric Maps (SPMs) revealed overlapping areas of visual and haptic encoding in adults and 7–8.5 year olds (Fig. 5A–E). Several whole brain contrasts were used during the sample encoding phase to indicate regions of activation for vision (VV+VH > rest; blue tones) and for haptics (HH+HV > rest; red tones). These contrasts are shown separately in the adults and in the 7–8.5-year-old children. As expected in adults, vision and haptics showed overlapping areas of activation bilaterally in the LOC, as well as in the IPS (IPS partially shown; Fig. 5A). Similarly in children, activation for vision and haptics overlapped bilaterally in the LOC and the IPS (Fig. 5B). Direct comparison of the two age



**Fig. 5.** Sample encoding phase: whole-brain overlap maps. Statistical Parametric Maps (SPMs) of group contrasts show overlapping areas of activation. (A) Overlap between vision (VV+VH > rest, balanced; navy blue) and haptics (HH+HV > rest, balanced; dark red) in adults. (B) Overlap between vision (VV+VH > rest, balanced; light blue) and haptics (HH+HV > rest, balanced; salmon) in 7–8.5 year olds. (C) Overlap between adults (dark red) and 7–8.5 year olds (light blue) for vision (VV+VH > rest, balanced). (D) Overlap between adults (navy blue) and 7–8.5 year olds (light blue) for haptics (HH+HV > rest, balanced). (E) Overlap between (A)–(D) for the LOC ROI selection (green arrow). On this and subsequent figures, colored lines on the sagittal plane correspond to axial slices along the z-axis and coronal slices along the y-axis. Functional data are presented at a threshold of  $p < 0.05$  (corrected) on an averaged T1-weighted anatomical image of all participants. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 6.** Sample encoding phase: whole-brain difference maps. SPMs of statistical differences between vision and haptics during the sample encoding phase in: (A) adults; and (B) 7–8.5 year olds. Regions showing significantly greater activity for haptics than vision ( $HH+HV > VV+VH$ ) is depicted in orange (adults) and yellow (7–8.5 year olds). Regions showing significantly greater activity for vision than haptics ( $VV+VH > HH+HV$ ) is depicted in blue (adults) and light blue (7–8.5 year olds). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

groups found no significant clusters, indicating very similar regions of neural activity during haptic sample encoding—the overlap between adults and 7–8.5 year olds encompassed bilateral areas in the LOC and the IPS (Fig. 5C). During visual sample encoding, both age groups showed bilateral activity in the LOC (Fig. 5D) and, similar to haptics, there were no significant differences between groups. In both haptic and visual sample encoding analyses, there was a trend for children’s activation to be more widespread than adults (see Supplementary Fig. S1).

### 3.1.2. Differences between maps of activation

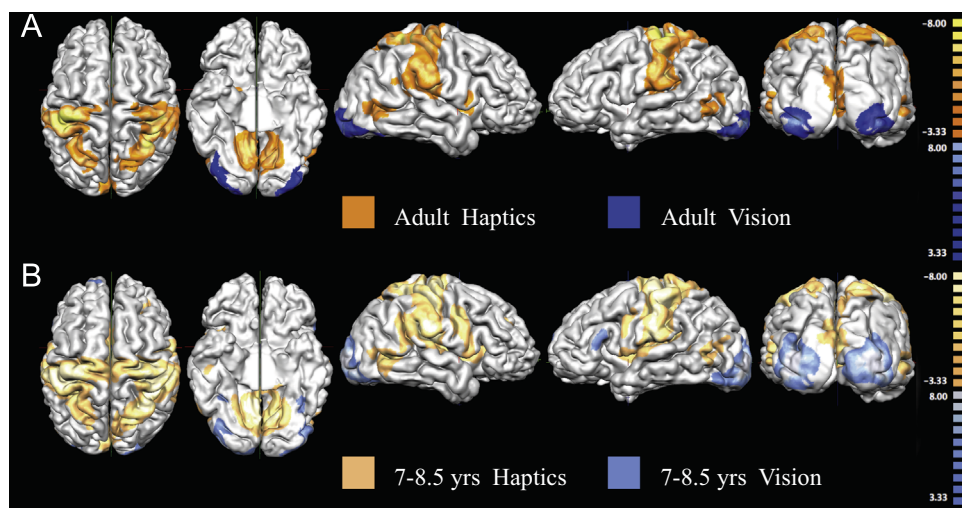
Complementary to the overlap analysis just described, whole-brain contrasts were also used to examine the differences between visual and haptic sample encoding in adults and in 7–8.5 year olds (Figs. 6–7). In these figures, regions of neural activity in which vision was greater than haptics ( $VV+VH > HH+HV$ ) are depicted in blue (adults) and light blue (7–8.5 year olds); regions in which haptics was greater than vision ( $HH+HV > VV+VH$ ) are depicted in orange (adults) and yellow (7–8.5 year olds). In adults, visual sample encoding activated areas of visual cortex bilaterally (Fig. 6A,  $Z = -4$ ), while haptic sample encoding activated bilateral areas in the IPS (Fig. 6A,  $Y = -67$ , top). Similar patterns of activation were found in 7–8.5 year olds (Fig. 6B). Critically, results

indicated bilateral middle temporal/occipital regions (MTG/MOG) in the putative visual cortex that showed significantly greater activity for haptics than for vision in both age groups (Fig. 6; see Fig. 7 for a 3D depiction, lateral views). These haptic-biased regions in MTG/MOG did not overlap with the bimodal regions described above, suggesting a haptic-preferring region in the visual stream that is not part of LOTv. Examining the overlap and difference maps together revealed a medial to lateral organization that transitioned from a visual bias in activation on the medial fusiform gyrus (FG) to a haptic bias in activation on the MTG/MOG, with the bimodal LOTv region in between on the lateral occipitotemporal sulcus (see Fig. 8A). This organization is illustrated in a gradient of visual to haptic biases during the sample encoding phase (displayed on “flooded” maps in Fig. 8B), which was used to localize the ROIs. Importantly, this pattern was present during the test matching phase (Fig. 8C).

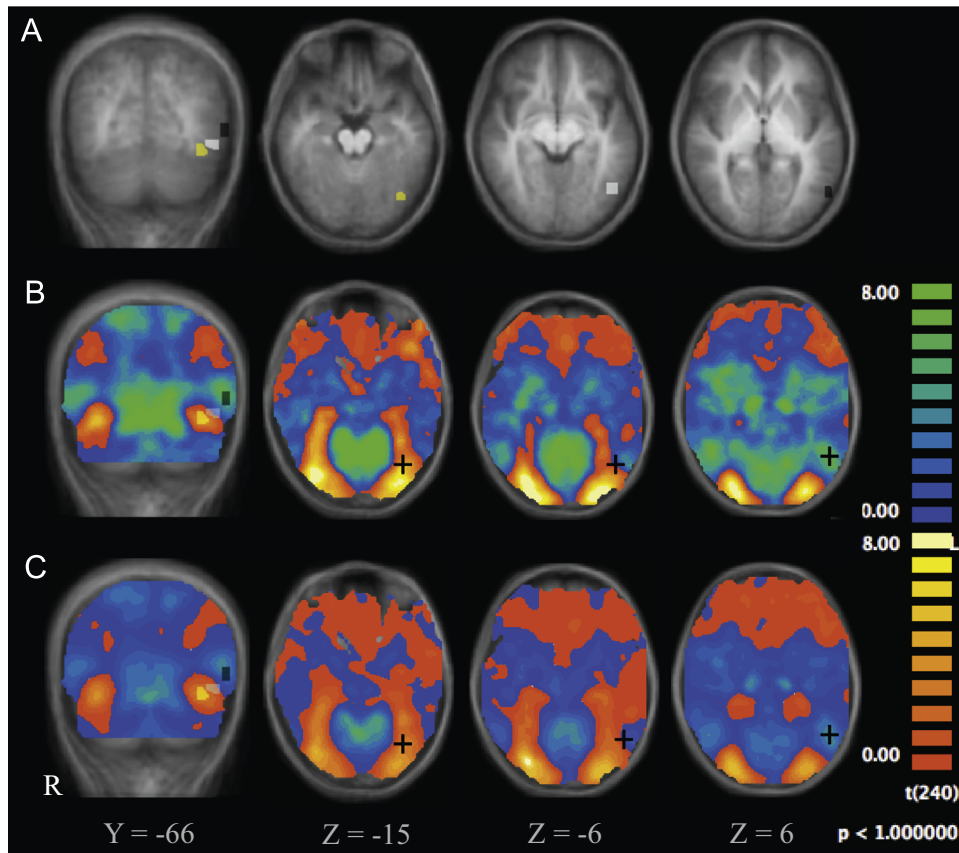
## 3.2. Test matching phase

### 3.2.1. Whole brain activation

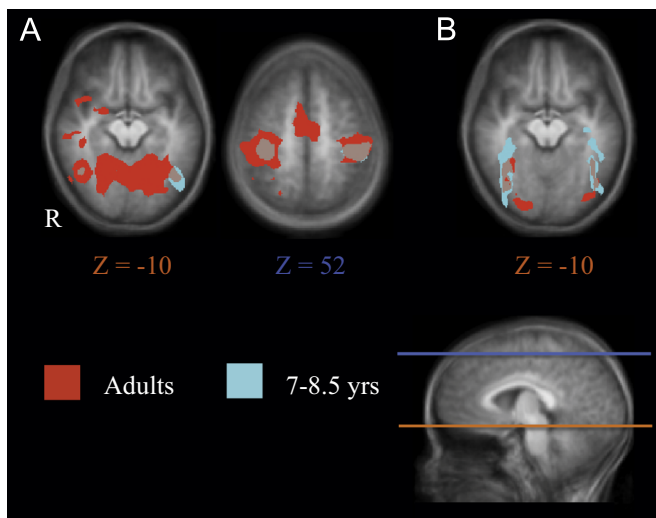
A balanced contrast comparing the experimental conditions against rest (i.e.,  $VV+HH+VH+HV > rest$ ) during the test matching phase also revealed known multisensory visuohaptic



**Fig. 7.** Sample encoding phase: overall difference. Statistically significant differences between vision and haptics during the sample encoding phase in: (A) adults; and (B) 7–8.5 year olds. Significant areas of activation for haptics ( $HH+HV > VV+VH$ ) are shown in orange (adults) and yellow (7–8.5 year olds), as compared to areas of activation for vision ( $VV+VH > HH+HV$ ), which are shown in blue (adults) and light blue (7–8.5 year olds). Views from left to right: superior, inferior, right hemisphere, left hemisphere, and posterior. Statistical maps are overlaid on a representative 3D anatomical image. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 8.** Medial to lateral transition. (A) Locations of the medial FG ROI (yellow), overlap LOC ROI (gray), and lateral MTG/MOG ROI (black). (B) Flooded SPMs ( $VV+VH > HH+HV$ ) show a gradient of visual (orange) to haptic (green) biases from medial to lateral regions during the sample encoding phase in all participants. (C) Flooded SPMs ( $VV+HV > HH+VH$ ) show a similar gradient during the test matching phase; crosshairs indicate ROI peaks. Crosshairs in B are for reference only. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 9.** Test matching phase: whole-brain maps. (A) SPMs of group contrasts between all conditions versus rest ( $VV+VH+HH+HV > \text{rest}$ , balanced). Crossmodal and intramodal matching in children and adults activated an overlapping network of regions that has been shown to be involved in multisensory visuohaptic object recognition. Regions include the LOC ( $Z = -10$ ) and the IPS ( $Z = 52$ ). (B) Crossmodal as compared to intramodal matching. SPM of the conjunction contrast for haptic-to-visual matching as compared to visual-only and haptic-only matching (i.e.,  $(HV > VV) \cap (HV > HH)$ ) activated the LOC bilaterally ( $Z = -10$ ) and overlapped between children and adults. Data from adults (red) and children (blue) are presented at a threshold of  $p < 0.05$  (corrected). LOC=lateral occipital complex; IPS=intraparietal sulcus. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

substrates in children and adults, including the LOC and the IPS (Fig. 9A; Amedi et al., 2001, 2002, 2005; Jao et al., 2014; James et al., 2002; James and Kim, 2010; Stilla and Sathian, 2008). Based on these whole-brain results, the network of neural substrates underlying visuohaptic processing appears to be recruited not only in adults, but importantly, are similarly activated in children by 7 years of age for this task.

To reveal specific effects of crossmodal enhancement (i.e., increased activation for crossmodal than intramodal sequentially-matched stimuli) throughout the brain, we compared crossmodal to intramodal matching using a conjunction of two contrasts, namely, the crossmodal haptic-to-visual matching task versus both intramodal tasks (i.e.,  $(HV > VV) \cap (HV > HH)$ ; Fig. 9B). This is similar to an intersection contrast in a PET study performed by Hadjikhani and Roland (1998), who discovered activation in the insula-claustrum only and suggested that this region must play a crucial role in the integration of crossmodal inputs. While we did not find activation in this region with our contrast (see Remedios et al. (2010) for evidence against the claustrum as an integrator of sensory information), we did find activation in the LOC bilaterally that overlapped across children and adults. The analogous intersection contrast with crossmodal visual-to-haptic matching (i.e.,  $(VH > VV) \cap (VH > HH)$ ) showed a less stable pattern. The only cluster was found in the post-central gyrus in adults, but a contrast across groups did not reveal a significant difference in this region.

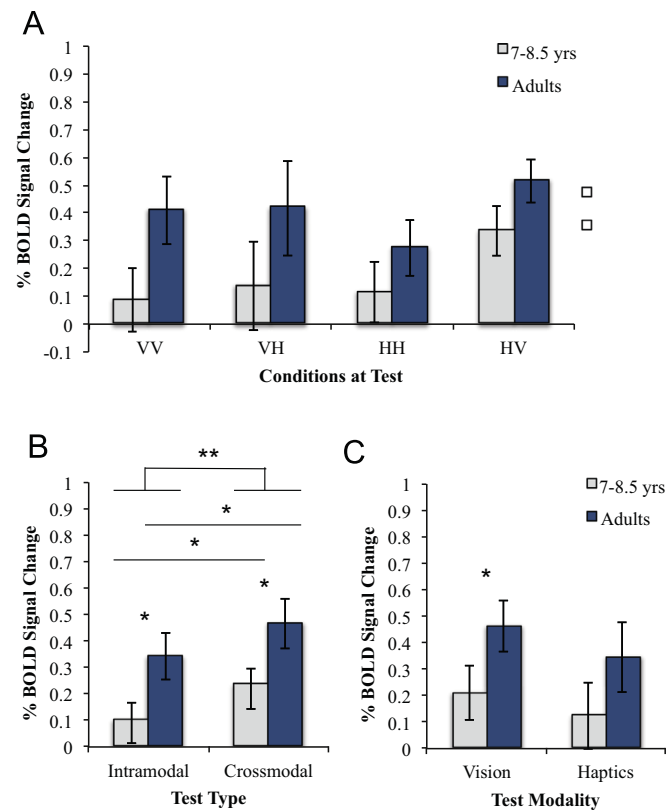
### 3.2.2. Region-of-interest results

To assess the pattern of activation during the test matching phase in the LOC ROI (see Fig. 5E, green arrow for overlapping contrasts used to define the ROI; Table 2), a repeated measures



**Table 2**  
Talairach coordinates (*x*, *y*, *z*), peak *t*-values, *p*-values, and number of voxels for the overlap, medial, and lateral ROIs at test.

Region	<i>x</i>	<i>y</i>	<i>z</i>	<i>t</i> -value	<i>p</i> -value	No. of voxels
Overlap: LOC	49	65	6	4.346	.001	756
Medial: FG	37	71	15	3.759	.001	618
Lateral: MTG/MOG	53	62	6	4.127	.001	578



**Fig. 10.** Test matching phase: LOC ROI. Percentage BOLD signal change from the group-based ROI is presented for each group of participants during the test matching phase for: (A) overall conditions; (B) test type (intramodal: VV and HH; crossmodal: VH and HV); and (C) test modality (vision: VV and HV; haptics: HH and VH).

(2 × 2) ANOVA was performed on the data extracted from that ROI (Fig. 10A) with test modality (visual or haptic) and test trial type (intramodal or crossmodal) as the within-subjects factors, and age group (7–8.5 year olds or adults) as the between-subjects factor. Results indicated a main effect of test trial type ( $F(1,19)=11.992$ ,  $p=.003$ ; Fig. 10B), with greater response for crossmodal than intramodal processing overall ( $t(20)=3.562$ ,  $p=.002$ ). This crossmodal enhancement effect was present in 7 year olds ( $t(10)=2.344$ ,  $p=.041$ ), as well as in adults ( $t(9)=2.680$ ,  $p=.025$ ). Furthermore, there was a main effect of group ( $F(1,19)=5.308$ ,  $p=.033$ ). Adults showed significantly higher BOLD signal change than 7–8.5 year olds for each test trial type (intramodal: ( $t(19)=2.191$ ,  $p=.041$ ); crossmodal: ( $t(19)=2.138$ ,  $p=.046$ )). There was no main effect of modality at test, however, suggesting fairly equivalent processing of both visual and haptic inputs in this region during the matching of objects for recognition (Fig. 10C). Interestingly, adults showed significantly higher BOLD signal change than 7–8.5 year olds during the visual test modality ( $t(19)=2.391$ ,  $p=.027$ ). This result supports previous findings in which the LOC becomes increasingly visually dominant with development, particularly for object-preference (Jao et al., 2014).

To examine the activation patterns in the medial FG and lateral

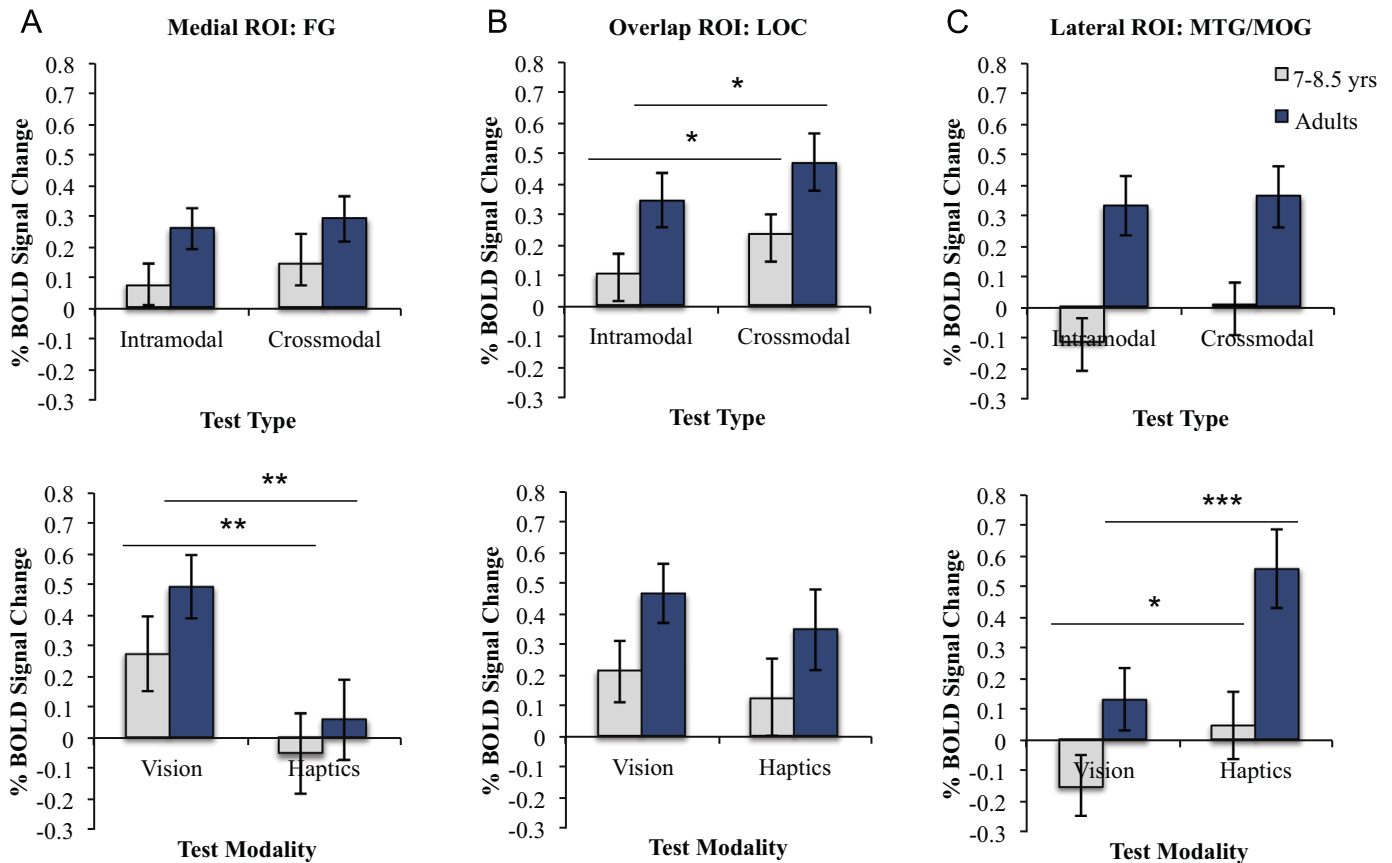
MTG/MOG ROIs, particularly in relation to the overlapping LOC ROI during the test matching phase (Fig. 8A and C; Table 2), a repeated measures (2 × 2 × 3) ANOVA was performed with age group (7–8.5 year olds or adults) as the between-subjects factor, and test modality (visual or haptic), test trial type (intramodal or crossmodal), and ROI (medial, overlap, lateral) as the within-subjects factors. Results showed a main effect of age group ( $F(1,19)=8.895$ ,  $p=.008$ ), with significantly higher BOLD signal change overall in adults than in 7–8.5 year olds ( $t(19)=2.982$ ,  $p=.008$ ), as well as a main effect of ROI ( $F(2,18)=7.398$ ,  $p=.005$ ). The statistical comparisons shown in Fig. 11 are within each age group and separated by region. Further results indicated a test type × ROI interaction effect ( $F(2,18)=4.148$ ,  $p=.033$ ); there was a greater response for crossmodal than intramodal processing within each group (7–8.5 year olds: ( $t(10)=2.344$ ,  $p=.041$ ); adults: ( $t(9)=2.680$ ,  $p=.025$ )) in the overlapping LOC ROI (Fig. 11B, top; also see Fig. 10B). This crossmodal enhancement effect, however, was not present in either age group in the medial FG ROI (Fig. 11A, top) or the lateral MTG/MOG ROI (Fig. 11C, top). Finally, there was a test modality × ROI interaction effect ( $F(2,18)=49.064$ ,  $p<.001$ ). In the medial ROI, BOLD responses were significantly higher when processing visual than haptic inputs in each group (7–8.5 year olds: ( $t(10)=4.264$ ,  $p=.002$ ); adults: ( $t(9)=4.330$ ,  $p=.002$ ); Fig. 11A, bottom). In the overlap ROI, there were no significant differences between visual and haptic processing (Fig. 11B, bottom; see also Fig. 10C). Lastly, in the lateral ROI, BOLD responses were higher when processing haptic than visual inputs in each group (7–8.5 year olds: ( $t(10)=2.536$ ,  $p=.030$ ); adults: ( $t(9)=6.213$ ,  $p<.001$ ); Fig. 11C, bottom).

#### 4. Discussion

The present study used functional MRI to investigate the development of crossmodal visuohaptic object recognition. The main finding was that visuohaptic crossmodal matching produced greater activation than intramodal matching in the LOC for both adults and children. To our knowledge this is the first study to find this effect in the LOC in adults or children. The key developmental finding was that children from 7–8.5 years of age did not differ qualitatively in terms of the overall pattern of activation, even though children did produce less activation in general across all conditions in the LOC during test matching. Additionally, the data in both children and adults revealed a haptic-preferring region in the bilateral middle temporal/occipital gyrus, a putative visual region, that was not considered “bimodal”. A broader perspective showed that the ventral occipitotemporal cortex followed a medial to lateral organization that transitioned from a visual to bimodal to haptic pattern of activation in both age groups. Although the LOC is a known region of visuohaptic convergence, these results provide novel insights into the mechanisms invoked for sequential information sharing across sensory modalities, as well as into the development of those mechanisms.

##### 4.1. Developmental similarities and differences

Both adults and children showed crossmodal enhancement effects, and recruited similar multisensory systems during visuohaptic object recognition. Neural activity in children, however, was generally more widespread during sample encoding, and weaker within a specialized bimodal area (i.e., the LOC) during test matching, than in adults. This suggests that although the neural mechanisms supporting crossmodal visuohaptic object processing are in place by 7 years, they are still undergoing change. According to psychophysical findings, children aged 8–10 years begin to appear adult-like in terms of integrating visual and haptic



**Fig. 11.** Test matching phase: Medial to lateral ROIs. Percentage BOLD signal change during the test matching phase is presented for each group of participants. Comparisons within test type (intramodal: VV and HH; crossmodal: VH and HV) and test modality (vision: VV and HV; haptics: HH and VH) are shown for the group-defined: (A) medial ROI (FG); (B) overlap ROI (LOC); and (C) lateral ROI (MTG/MOG).

information for certain aspects of form discrimination (e.g., size and orientation; Gori et al., 2008). Prior to this age, however, children were shown to be unable to combine sensory information for perception in an “optimal” manner. These behavioral results suggest that there is a transitional stage of development prior to 8 years during which visuohaptic integration for specific types of shape discrimination becomes optimized. There is also prior neural evidence that the LOC continues to develop after 7 years of age with continued experience on some visual or visuohaptic tasks. For instance, previous work has shown that visual dominance in the LOC during visuohaptic object recognition does not reach adult levels until after 8.5 years (Jao et al., 2014). More specifically, it was found that visual dominance in the LOC continued to increase from 8.5 years of age into young adulthood, while activation for the haptic modality remained fairly constant from 4 years into adulthood. This finding was replicated in the current study in which BOLD activation for vision was higher in adults than in children during the test matching phase.

The age ranges for the current study were therefore selected with the expectation that the fMRI data would follow the patterns seen in previous studies showing a transitional, non-adult-like level of multisensory processing in children. This was indeed the case, as children not only showed more widespread whole-brain activity than adults during sample encoding, they also showed lower levels of BOLD activation in the LOC during test matching for crossmodal and intramodal processing. The selection was also based to some degree on the ability of children at different ages to perform the delayed match-to-sample tasks successfully, which required the participants to maintain the sample stimulus in memory during each of the subsequent matching trials. In the current study, all of the 7–8.5-year-old children were able to do so

successfully. While testing an even younger age range may provide further details about the developmental trajectory of visuohaptic integration, the current findings are interpretable in their own right.

#### 4.2. Haptic-preferring regions within the LOC

Contrasts of vision versus haptics revealed robust effects in both age groups in which bilateral regions in the occipitotemporal cortex showed greater activity for haptics than for vision. While other adult studies have found haptic object-selectivity—often defined by a contrast of haptic shape versus haptic texture—in the LOC (Amedi et al., 2001; Lacey et al., 2010; Stilla and Sathian, 2008; Zhang et al., 2004; Stoesz et al., 2003; Prather et al., 2004; Peltier et al., 2007; for a review, see Lacey et al. (2009)), it has always been found to be overlapping with visually object-selective regions (Amedi et al., 2001; James et al., 2002; Stilla and Sathian 2008; Peltier et al., 2007; among others). Thus, ours is the first study to our knowledge that has reported a region in the ventral “visual” stream that responds significantly more during haptic than visual processing in both children and adults. This may be in part due to the majority of the aforementioned studies emphasizing the overlap of haptics and vision, rather than the difference between the two sensory modalities. There have, however, been indications of haptic activation in MTG/MOG in previous adult studies. One such study showed that somatosensory shape processing not only activated a region of the occipitotemporal cortex that overlapped a subregion of the visual LOC, namely in the LOTv, it also activated regions abutting the LOTv (Amedi et al., 2001); another showed activity in the MOG during visuotactile versus other bisensory object-related processing (Kassuba et al., 2011).

The current findings may not simply be the result of contrasting vision and haptics; they may be due to the context of the encoding phase, which was the sample phase of a crossmodal delayed match-to-sample task. Within the context of crossmodal object recognition, children and adults may be encoding the sample in such a way that haptic processing activates an additional region within the ventral occipitotemporal cortex that is driven more by haptics than by vision, although this region may not necessarily be object-selective. This hypothesis seems even more likely given that the participants knew beforehand whether the sample stimulus was to be matched to a crossmodal or intramodal target. Although this was an unexpected finding, it warrants further investigation, especially as it relates to the malleability of putative visual cortical signals by task demands.

Examining the whole of the ventral occipitotemporal cortex, the overlapping patterns of visual and haptic activation during test matching showed a medial to lateral organization that transitioned from a visual to haptic bias in the mapping of the sensory modalities used to process shape information. The mapping transitioned from regions of visual preference that were located more medially to regions of haptic preference that were located more laterally, with bimodal visuohaptic preference in the overlapping middle regions. Although not widely reported in the literature, there has been mention of a similar trend in at least one previous study. Specifically, whole-brain results of haptic and visual shape-selectivity showed the recruitment of bilateral regions that were each adjacent to the overlapping bimodal area in the LOC and opposite to one another, thus following a similar pattern; however, the transition from visual to haptic sensory preference was not addressed in terms of cortical organization or development (Stilla and Sathian, 2008). The finding of a haptic-preferring region along with visual regions in the LOC suggest that the multisensory signals integrated in the LOTv likely arrive via neighboring modality-biased regions within the LOC itself. We hypothesize that these modality-biased regions transform somatosensory and visual signals to facilitate multisensory processing in the LOTv. A future step should be identifying the exact nature of those transformations.

#### 4.3. Crossmodal effects in the LOC

##### 4.3.1. Crossmodal enhancement over intramodal matching

Crossmodal haptic-to-visual matching, which required the subject to compare visual test stimuli with the haptic sample, activated the LOC—a putative visual region—more strongly than the intramodal control conditions. The region of LOC that showed this effect overlapped in children and adults. Thus, in accordance with our first prediction, there was a stronger response for crossmodal than intramodal matching that was apparent in both age groups. This effect, however, was present only for haptic-to-visual crossmodal matching; contrasting visual-to-haptic crossmodal matching with intramodal conditions did not yield a significant effect in the LOC. This asymmetry is discussed in the next subsection (4.3.2 Crossmodal asymmetry). The ROI analysis revealed that, within each type of condition, adults showed higher levels of BOLD activity than 7 to 8.5 year olds. Thus, both intramodal and crossmodal processing are still developing in children, which reflects a decreased ability to process visual and haptic information about object shape. Additionally, the crossmodal enhancement effect was present in the overlapping bimodal LOC region (LOTv), but not in the medial visual-preferring or lateral haptic-preferring areas of the LOC. Together, these results suggest that the LOC not only processes crossmodal information, but also more importantly, is sensitive to sequential *changes* in sensory modality during object recognition.

One possible explanation for the crossmodal enhancement

effect is the reactivation of the encoded object at test that occurs in parallel with the activation of the test object. The LOC, as a multisensory region concerned with determining object shape, processes visual and haptic inputs in parallel. In situations in which information from one modality or another is unavailable, it must be able to detect any changes in sensory input and share the accessible information between modalities efficiently. Much of the research conducted on visuohaptic processing has now established that the LOC is bimodal in terms of its representations of visual and haptic (familiar and novel) shape information (Amedi et al., 2001; James et al., 2002; James, James, Humphrey et al., 2006; James and Kim, 2010; Lacey et al., 2010, 2014; Peltier et al., 2007; Pietrini et al., 2004; Stilla and Sathian, 2008; Stoesz et al., 2003; Zhang et al., 2004), and the current data also support this notion. Moreover, a previous study has demonstrated neural convergence of visual and haptic inputs in the LOC through inverse effectiveness (Kim et al., 2012). Based on these findings, it is plausible that during crossmodal matching, some of the population of neurons within this region would be reactivated at test (see Lacey and Sathian (2014) for a review of mental imagery), while others would be activated by the sensory percept. The combination of activated and re-activated neural populations would produce greater activation in crossmodal matching tasks, which require reactivation of the encoded stimulus as well as activation for the current sensory input, than in intramodal matching tasks, which do not.

Alternatively, it is possible that the crossmodal tasks may simply have been more difficult and required more intense processing than the intramodal matching tasks. While this is a possibility in the current study and should be addressed in future studies, the presence of a weak asymmetry in activation between the two crossmodal conditions seems to indicate that difficulty alone cannot explain the differences.

##### 4.3.2. Crossmodal asymmetry

While the present findings indicated crossmodal enhancement effects in the LOC, direct comparisons of crossmodal conditions (i.e., VH versus HV) during the test matching phase did not provide strong evidence for an asymmetry, or two-way directionality effect, of crossmodal processing. Comparisons between the HV and VH conditions indicated no significant differences in children or adults.

Although asymmetries within crossmodal processing have been addressed in previous studies, these have implemented different types of stimuli and have resulted in varying conclusions. In one particular study, Positron Emission Tomography (PET) was implemented to examine the effects of the presentation order of crossmodal information (Kawashima et al., 2002). Using crossmodal discrimination tasks in conditions analogous to the current study (i.e., visual-to-haptic, haptic-to-visual), their findings showed asymmetric processing of crossmodal stimuli (i.e., digital cylinders) such that only the visual-to-haptic presentation order activated the inferior temporal cortex, while both presentation orders activated the inferior parietal cortex. It was proposed that this was evidence for two different pathways underlying crossmodal discrimination depending on the temporal order of stimulus presentation (Kawashima et al., 2002). Similar types of directionality effects have been demonstrated recently in bilateral regions of the lateral occipital cortex and the aIPS using fMRI, but have only been found in one direction (Kassuba et al., 2013). This asymmetry during crossmodal matching of highly familiar objects occurred only in congruent visual-to-haptic conditions, with little effect of crossmodal matching on brain activation in these regions during haptic-to-visual or incongruent conditions. Based on these results, Kassuba and colleagues (2013) concluded that there is a modality-specific asymmetry with a preference supporting the

functional primacy of vision during visuohaptic object recognition. By contrast, a previous delayed match-to-sample study showed crossmodal enhancement in the left aIPS that was independent of matching direction (i.e., visual-to-haptic and haptic-to-visual; Grefkes et al., 2002). This particular study, however, used novel objects, as did the current study. The discrepant findings of asymmetry suggest that the types of stimuli used, namely novel versus familiar objects, may contribute to differences in crossmodal enhancement (Kassuba et al., 2013). That is, crossmodal enhancement may rely heavily on the specific experimental context wherein greater enhancement might be expected for novel than familiar objects (for a more in-depth discussion of these implications, see Kassuba et al. (2013)).

In the present study, whole-brain results from the test matching phase indicated that haptic-to-visual crossmodal recognition activated the LOC in both adults and children, while visual-to-haptic recognition had little effect in either age group. To examine this further, direct comparisons were made in the post-hoc ROI analysis, but results indicated no significant differences between the HV and VH conditions. As such, while it is possible that crossmodal processing in the LOC may be constrained by the dominant modality that provides the most salient or reliable of information at the time of recognition (e.g., functional primacy of vision)—which would be consistent with models of optimal integration where each sensory modality is weighted according to its reliability (Ernst and Banks, 2002; Helbig and Ernst, 2007)—our findings are not strong enough to support (or refute) this pattern of asymmetry. Nevertheless, the current findings lend support for the notion that the recruitment and specific location of multi-sensory convergence areas is highly contingent upon two primary factors: the information content being processed (e.g., shape), and the modality being used to process the crossmodal input during recognition (Amedi et al., 2005).

Overall, the current results and findings from previous studies seem to indicate that multisensory convergence cannot be described as a unitary developmental process, but instead should be described as several processes that follow differing developmental timelines. Some processes such as crossmodal recognition may require more experience in order to be optimized, and thus take longer to fully develop than others depending on the input. In terms of visual processing, prior behavioral evidence has shown that even children from 6 to 8 years have difficulty recognizing objects from unusual views (Bova et al., 2007; Juttner et al., 2006; Mondloch et al., 2003; Mondloch et al., 2002; see Nishimura et al. (2009) for a detailed review), and are relatively poor compared to adults at recognizing complex forms such as faces (Mondloch et al., 2006; although see Crookes and McKone (2009)). These delays in specific visual object recognition proficiencies implicate a delayed development of the occipitotemporal cortex, particularly of the LOC, but perhaps only for more specific types of shape processing. Thus, the processing of complex objects or abstract forms of inputs (e.g., crossmodal stimuli) may result in protracted developmental trajectories as compared to the processing of simple shapes. Studying the development of these processes constitutes grounds for future research.

#### 4.4. Handedness and the LOC

Bimanual exploration, which was implemented in the current study, may obviate the need to control for handedness. Previous fMRI findings in adults comparing left- and right-handed object exploration have shown bilateral LOtv activity that was not influenced by the hand-in-use (Amedi et al., 2010). Moreover, results from further adult studies have shown a left-lateralized bias in activation with either left-handed exploration (James, Servos, Kilgour, et al., 2006; Kilgour et al., 2005), right-handed exploration

(Kim et al., 2012), or bimanual exploration (Kim and James, 2010). Yet, it is important to note more recent findings suggesting that handedness may play a role in LOC activation (Yalachkov et al., 2015). In particular, one study demonstrated that the left LOC was activated more strongly by bimodal than unimodal stimuli when explored with the non-dominant (left) hand (Yalachkov et al., 2015). This suggests that the neural signal in the LOC during visuohaptic processing may depend on the hand-in-use.

Based on this body of evidence, it is arguable whether the hand-in-use or general handedness contributes to the neural response in higher cortical areas during visuohaptic object exploration. There is a possibility that the left-lateralized bias found in the LOC in some studies (e.g., Kim et al., 2012; Yalachkov et al., 2015) may be due not solely to the hand used to explore the stimulus, but also to the strong right-handed preference of the participants tested in these studies. Thus, to mitigate hand-in-use effects in the current study, haptic exploration was performed bimanually similar to other developmental studies of young children (Bushnell and Baxt, 1999; Kalagher and Jones, 2011a, 2011b). Additionally, to minimize differences in handedness preference, we recruited mostly right-handed participants, and did not include any strongly left-preferring participants. Clearly, future studies in children and adults could examine neural differences due to left- versus right- versus two-hand exploration in left- and right-handers during haptic and visuohaptic object recognition.

## 5. Conclusions

In summary, using a crossmodal delayed match-to-sample task with novel objects, we have investigated the neural substrates involved in visuohaptic processing in children and adults. Importantly, we argued that there is crossmodal enhancement in the LOC, suggesting that this region is sensitive to *changes* in sensory modality. We have shown that this effect, as well as the network of multi-sensory regions consistently found in adults, is present in children, although children show generally more widespread activity during sample encoding and weaker BOLD signal change in the LOC during test matching than adults. Finally, we have found evidence of a bilateral region in the MTG/MOG that is haptic-preferring in both age groups. This region abuts the bimodal LOtv, and importantly, indicates a medial to lateral organization with a visual to haptic transitional bias in the LOC that develops by 7 years of age.

## Acknowledgments

The authors would first and foremost like to thank all of the participants and parents for making this research possible. We wish to acknowledge Alyssa Kersey, Emily Thomas, Rachel Winchell, Rachel Crum, and JeanneMarie Heeb, as well as the members of the Cognition and Action Neuroimaging Lab at Indiana University, for the recruitment of subjects and collection of data. Thanks to Colleen McCracken and Sean Berry, the MR technicians, and Dr. Hu Cheng, the MR physicist, at the Indiana University Imaging Research Facility. This research was supported by: NSF, USA 0903495 (RJJ); NIH/NICHHD, USA HD057077 (KHJ) and 5T32HD007475 (KHJ); and the ME-TACyT Initiative of the Lilly Endowment Inc.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.neuropsychologia.2015.08.008>.

## References

- Amedi, A., Malach, R., Hendler, T., Peled, S., Zohary, E., 2001. Visuo-haptic object-related activation in the ventral visual pathway. *Nat. Neurosci.* 4, 324–330.
- Amedi, A., Jacobson, G., Hendler, T., Malach, R., Zohary, E., 2002. Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cereb. Cortex* 12, 1202–1212.
- Amedi, A., Raz, N., Azulay, H., Malach, R., Zohary, E., 2010. Cortical activity during tactile exploration of objects in blind and sighted humans. *Restor. Neurol. Neurosci.* 28, 143–156.
- Amedi, A., von Kriegstein, K., van Atteveldt, N.M., Beauchamp, M.S., Naumer, M.J., 2005. Functional imaging of human crossmodal identification and object recognition. *Exp. Brain Res.* 166, 559–571.
- Binkofski, F., Buccino, G., Stephan, K.M., Rizzolatti, G., Seitz, R.J., Freund, H.J., 1999. A parieto-premotor network for object manipulation: evidence from neuroimaging. *Exp. Brain Res.* 128, 210–213.
- Bodegard, A., Geyer, S., Grefkes, C., Zilles, K., Roland, P.E., 2001. Hierarchical processing of tactile shape in the human brain. *Neuron* 31, 317–328.
- Bova, S.M., Fazzi, E., Giovenzana, A., Montomoli, C., Signorini, S.G., Zoppello, M., Lanzi, G., 2007. The development of visual object recognition in school-age children. *Dev. Neuropsychol.* 31, 79–102.
- Bushnell, E.W., Baxt, C., 1999. Children's haptic and cross-modal recognition with familiar and unfamiliar objects. *J. Dev. Psychol.: Hum. Percept. Perform.* 25, 1867–1881.
- Bushnell, E.W., Boudreau, J.P., 1993. Motor development and the mind: the potential role of motor abilities as a determinant of aspects of perceptual development. *Child Dev.* 64, 1005–1021.
- Crookes, K., McKone, E., 2009. Early maturity of face recognition: no childhood development of holistic processing, novel face encoding, or face-space. *Cognition* 111 (2), 219–247.
- Culham, J.C., Kanwisher, N.G., 2001. Neuroimaging of cognitive functions in human parietal cortex. *Curr. Opin. Neurobiol.* 11, 157–163.
- Deibert, E., Kraut, M., Kremen, S., Hart Jr., J., 1999. Neural pathways in tactile object recognition. *Neurology* 52, 1413–1417.
- Easton, R.D., Greene, A.J., Srinivas, K., 1997. Transfer between vision and haptics: memory for 2-D patterns and 3-D objects. *Psychon. Bull. Rev.* 4, 403–410.
- Ernst, M.O., Banks, M.S., 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.
- Gentile, G., Petkova, V.I., Ehrsson, H.H., 2011. Integration of visual and tactile signals from the hand in the human brain: an fMRI Study. *J. Neurophysiol.* 105, 910–922.
- Gori, M., Del Viva, M., Sandini, G., Burr, D.C., 2008. Young children do not integrate visual and haptic form information. *Curr. Biol.* 18, 694–698.
- Grefkes, C., Weiss, P.H., Zilles, K., Fink, G.R., 2002. Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. *Neuron* 35, 173–184.
- Grill-Spector, K., Golarai, G., Gabrieli, J., 2008. Developmental neuroimaging of the human ventral visual cortex. *Trends Cogn. Sci.* 12, 152–162.
- Grill-Spector, K., Kourtzi, Z., Kanwisher, N., 2001. The lateral occipital complex and its role in object recognition. *Vis. Res.* 41, 1409–1422.
- Hadjikhani, N., Roland, P.E., 1998. Cross-modal transfer of information between the tactile and the visual representations in the human brain: a positron emission tomographic study. *J. Neurosci.* 18, 1072–1084.
- Helbig, H.B., Ernst, M.O., 2007. Optimal integration of shape information from vision 803 and touch. *Exp. Brain Res.* 179, 595–606.
- James, K.H., Swain, S., Jones, S.S., Smith, L.B., 2014. Young children's self-generated object views and object recognition. *J. Cogn. Dev.* 15 (3), 393–401.
- James, T.W., Humphrey, G.K., Gati, J.S., Servos, P., Menon, R.S., Goodale, M.A., 2002. Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia* 40, 1706–1714.
- James, T.W., James, K.H., 2013. Expert individuation of objects increases activation in the fusiform face area of children. *NeuroImage* 67, 182–192.
- James, T.W., James, K.H., Humphrey, G.K., Goodale, M.A., 2006. Do visual and tactile object representations share the same neural substrate?. In: Heller, M.A., Balasteros, S. (Eds.), *Touch and Blindness: Psychology and Neuroscience*. Lawrence Erlbaum, Mahwah, NJ.
- James, T.W., Kim, S., 2010. Dorsal and ventral cortical pathways for visuo-haptic shape integration revealed using fMRI. In: Naumer, M.J., Kaiser, J. (Eds.), *Multisensory Object Perception in the Primate Brain*. Springer, New York.
- James, T.W., Kim, S., Fisher, J.S., 2007. The neural basis of haptic object processing. *Can. J. Exp. Psychol.* 61, 219–229.
- James, T.W., Servos, P., Kilgour, A.R., Huh, E., Lederman, S., 2006. The influence of familiarity on brain activation during haptic exploration of 3-D facemasks. *Neurosci. Lett.* 397, 269–273.
- Jao, R.J., James, T.W., James, K.H., 2014. Multisensory convergence of visual and haptic object preference across development. *Neuropsychologia* 56, 381–392.
- Juttner, M., Muller, A., Rentschler, I., 2006. A developmental dissociation of view-dependent and view-invariant object recognition in adolescence. *Behav. Brain Res.* 175, 420–424.
- Kalagher, H., Jones, S.S., 2011a. Developmental change in young children's use of haptic information in a visual task: the role of hand movements. *J. Exp. Child Psychol.* 108, 293–307.
- Kalagher, H., Jones, S.S., 2011b. Young children's haptic exploratory procedures. *J. Exp. Child Psychol.* 110, 592–602.
- Kassuba, T., Klinge, C., Holig, C., Menz, M.M., Pttio, M., Roder, B., Siebner, H.R., 2011. The left fusiform gyrus hosts trisensory representations of manipulable objects. *NeuroImage* 56, 1566–1577.
- Kassuba, T., Klinge, C., Holig, C., Roder, B., Siebner, H.R., 2013. Vision holds a greater share in visuo-haptic object recognition than touch. *NeuroImage* 65, 59–68.
- Kawashima, R., Watanabe, J., Kato, R., Nakamura, A., Hatano, K., Schormann, T., Sato, K., Fukuda, H., Ito, K., Zilles, K., 2002. Direction of cross-modal information transfer affects human brain activation: a PET study. *Eur. J. Neurosci.* 16, 137–144.
- Kilgour, A.R., Kitada, R., Servos, P., James, T.W., Lederman, S.J., 2005. Haptic face identification activates ventral occipital and temporal areas: an fMRI study. *Brain Cogn.* 59, 246–257.
- Kim, S., James, T.W., 2010. Enhanced effectiveness in visual-haptic object-selective brain regions with increasing stimulus salience. *Hum. Brain Mapp.* 31, 678–693.
- Kim, S., Stevenson, R.A., James, T.W., 2012. Visuo-haptic neuronal convergence demonstrated with an inversely effective pattern of BOLD activation. *J. Cogn. Neurosci.* 24 (4), 830–842.
- Klatzky, R.L., Lederman, S.J., Reed, C.L., 1987. There's more to touch than meets the eye: The salience of object attributes for haptics with and without vision. *J. Exp. Psychol.: Gen.* 116, 356–369.
- Kourtzi, Z., Kanwisher, N., 2001. Representation of perceived object shape by the human lateral occipital complex. *Science* 293, 1506–1509.
- Lacey, S., Flueckiger, P., Stilla, R., Lava, M., Sathian, K., 2010. Object familiarity modulates the relationship between visual object imagery and haptic shape perception. *NeuroImage* 49, 1977–1990.
- Lacey, S., Sathian, K., 2014. Visuo-haptic multisensory object recognition, categorization, and representation. *Front. Psychol.* 5 (730), 1–15.
- Lacey, S., Sathian, K., 2011. Multisensory object representation: Insights from studies of vision and touch. *Prog. Brain Res.* 191, 165–176.
- Lacey, S., Stilla, R., Sreenivasan, K., Deshpande, G., Sathian, K., 2014. Spatial imagery in haptic shape perception. *Neuropsychologia* 60, 144–158.
- Lacey, S., Tal, N., Amedi, A., Sathian, K., 2009. A putative model of multisensory object representation. *Brain Topogr.* 21, 269–274.
- Lederman, S.J., Klatzky, R.L., 1993. Extracting object properties through haptic exploration. *Acta Psychol.* 84, 29–40.
- Lederman, S.J., Klatzky, R.L., 1990. Haptic classification of common objects: knowledge-driven exploration. *Cogn. Psychol.* 22, 421–459.
- Lederman, S.J., Klatzky, R.L., 1987. Hand movements: a window into haptic object recognition. *Cogn. Psychol.* 19, 342–368.
- Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. USA* 92, 8135–8139.
- Mondloch, C.J., Geldart, S., Maurer, D., Le Grand, R., 2003. Developmental changes in face processing skills. *J. Exp. Child Psychol.* 86, 67–84.
- Mondloch, C.J., Le Grand, R., Maurer, D., 2002. Configural face processing develops more slowly than featural face processing. *Perception* 31, 553–566.
- Mondloch, C.J., Maurer, D., Ahola, S., 2006. Becoming a face expert. *Psychol. Sci.* 17 (11), 930–934.
- Newell, F.N., Ernst, M.O., Tjan, B.S., Bulthoff, H.H., 2001. Viewpoint dependence in visual and haptic object recognition. *Psychol. Sci.* 12, 37–42.
- Nishimura, M., Scherf, S., Behrmann, M., 2009. Development of object recognition in humans. *Biol. Rep.* 1 (56), 1–4.
- Norman, J.F., Clayton, A.M., Norman, H.F., Crabtree, C.E., 2008. Learning to perceive differences in solid shape through vision and touch. *Perception* 37, 185–196.
- Norman, J.F., Norman, H.F., Clayton, A.M., Lianekhammy, J., Zielke, G., 2004. The visual and haptic perception of natural object shape. *Percept. Psychophys.* 66, 342–351.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Peltier, S., Stilla, R., Mariola, E., LaConte, S., Hu, X., Sathian, K., 2007. Activity and effective connectivity of parietal and occipital cortical regions during haptic shape perception. *Neuropsychologia* 45, 476–483.
- Pietrini, P., Furey, M.L., Ricciardi, E., Gobbini, M.I., Wu, W.H., Cohen, L., Guazzelli, M., Haxby, J.V., 2004. Beyond sensory images: Object-based representation in the human ventral pathway. *Proc. Natl. Acad. Sci. USA* 101, 5658–5663.
- Prather, S.C., Votaw, J.R., Sathian, K., 2004. Task-specific recruitment of dorsal and ventral visual areas during tactile perception. *Neuropsychologia* 42, 1079–1087.
- Reed, C.L., Shoham, S., Halgren, E., 2004. Neural substrates of tactile object recognition: an fMRI study. *Hum. Brain Mapp.* 21, 236–246.
- Remedios, R., Logothetis, N.K., Kayser, C., 2010. Unimodal responses prevail within the multisensory claustrum. *J. Neurosci.* 30 (39), 12902–12907.
- Roland, P.E., O'Sullivan, B., Kawashima, R., 1998. Shape and roughness activate different somatosensory areas in the human brain. *Proc. Natl. Acad. Sci. USA* 95, 3295–3300.
- Ruff, H.A., 1984. Infants' manipulative exploration of objects: effects of age and object characteristics. *Dev. Psychol.* 20, 9–20.
- Ruff, H.A., 1986. Components of attention during infants' manipulative exploration. *Child Dev.* 57, 105–114.
- Ruff, H.A., 1989. The infant's use of visual and haptic information in the perception and recognition of objects. *Can. J. Psychol.* 43, 302–319.
- Saito, D.N., Okada, T., Morita, Y., Yonekura, Y., Sadato, N., 2003. Tactile-visual cross-modal shape matching: a functional MRI study. *Cogn. Brain Res.* 17, 14–25.
- Scherf, K.S., Behrmann, M., Humphreys, K., Luna, B., 2007. Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Dev. Sci.* 10 (4), F15–F30.
- Stein, B.E., Stanford, T.R., 2008. Multisensory integration: current issues from the

- perspective of the single neuron. *Nat. Rev. Neurosci.* 9, 255–266.
- Stilla, R., Sathian, K., 2008. Selective visuo-haptic processing of shape and texture. *Hum. Brain Mapp.* 29, 1123–1138.
- Stoesz, M.R., Zhang, M., Weisser, V.D., Prather, S.C., Mao, H., Sathian, K., 2003. Neural networks active during tactile form perception: common and differential activity during macrospatial and microspatial tasks. *Int. J. Psychophysiol.* 50, 41–49.
- Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system – an approach to cerebral imaging. Thieme Medical Publishers, New York.
- Tootell, R.B.H., Dale, A.M., Sereno, I., Malach, R., 1996. New images from the human visual cortex. *Trends Neurosci.* 19, 481–489.
- Wakefield, E.M., James, T.W., James, K.H., 2013. Neural correlates of gesture processing across human development. *Cogn. Neuropsychol.* 30, 58–76.
- Yalachkov, Y., Kaiser, J., Doehrmann, O., Naumer, M.J., 2015. Enhanced visuo-haptic integration for the non-dominant hand. *Brain Res.* 1614, 75–85.
- Zhang, M., Weisser, V.D., Stilla, R., Prather, S.C., Sathian, K., 2004. Multisensory cortical processing of object shape and its relation to mental imagery. *Cogn. Affect. Behav. Neurosci.* 4, 251–259.