Chapter 13

Dorsal and Ventral Cortical Pathways for Visuo-haptic Shape Integration Revealed Using fMRI

Thomas W. James and Sunah Kim

13.1 Introduction

Visual object recognition is pervasive and central to many aspects of human functioning. In adults, it seems effortless and nearly automatic. Despite the ease with which we perceive and identify objects, however, computer simulations of object recognition have been largely unsuccessful at mimicking human recognition. Simulations can succeed in constrained environments, but cannot match the flexibility of the human system. One reason machine vision may have had limited success outside of highly constrained contexts is that visual recognition is an extremely difficult computational problem (Lennie, 1998). Another reason, however, may be that computational approaches have largely restricted themselves to modeling the visual system in isolation from other sensory and motor systems, whereas human visual recognition is embedded in interactions between multiple sensory systems (Clark, 1997; de Sa and Ballard, 1998). Although research of multisensory phenomena has a long history (Molyneux, 1688), research into the neural mechanisms of sensory processes in humans and other primates has been dominated in recent years by investigations of unisensory visual function. This has led to a relative paucity of empirical data from – and theoretical treatment of – other sensory systems and, perhaps most importantly, interactions between multiple sensory systems. Our goals in this chapter are twofold. First, we describe an influential theoretical perspective on the organization of the cortical visual system, the two visual streams theory, and apply that perspective to interactions between visual and haptic object shape processing. Second, using a new methodology, we assess neuronal convergence of visual and haptic inputs in regions considered part of those two separable pathways.

T.W. James (✉)
Department of Psychological and Brain Sciences, Cognitive Science Program, Indiana University, 1101 E Tenth St., Bloomington, IN 47405, USA
e-mail: thwjames@indiana.edu

M.J. Naumer, J. Kaiser (eds.), *Multisensory Object Perception in the Primate Brain*,
DOI 10.1007/978-1-4419-5615-6_13, © Springer Science+Business Media, LLC 2010

13.2 Visual Cortical Pathways for Action and Perception

For almost three decades, one of the most dominant theories of visual system organization has been the two visual streams theory (Goodale and Milner, 1992; Ungerleider and Mishkin, 1982). There are two prominent sets of visual projections in the primate cerebral cortex: the ventral stream, which arises in area V1 and projects to the inferotemporal cortex, and the dorsal stream, which also arises in area V1 and projects to the posterior parietal cortex (Baizer et al., 1991; Morel and Bullier 1990; Ungerleider and Mishkin, 1982; Young, 1992). In the early 1990s, Goodale and Milner (Goodale and Milner, 1992; Milner and Goodale, 1995) argued that the ventral stream plays the major role in constructing our perceptual representation of the visual world and the objects within it, while the dorsal stream mediates the visual control of actions that we direct at those objects. This idea was different from the initial conceptualization of "what–where" dual pathways (Ungerleider and Mishkin, 1982). The what–where hypothesis suggested that both streams were involved in perception, but different aspects of perception. The ventral (what) stream was intimately involved in identifying objects, whereas the dorsal (where) stream was involved in locating them in space.

In the Goodale and Milner model, the same information (that is, the same rudimentary features of objects) is processed by both streams, but for different purposes. In other words, the input to both streams is the same, but the outputs are different. Because the inputs are the same and the outputs are different, the calculations and transformations that each stream performs on the input must be different. In the case of the ventral stream, object features and parameters are transformed in such a way that it produces our phenomenological experience of the world, allowing us to deliberate and reason about our choice of actions. In the case of the dorsal stream, the same object information is transformed in such a way that it is useful for *controlling* those actions.

Some of the most compelling evidence for Goodale and Milner's perception–action hypothesis has come from studies of patient DF, a young woman who suffered irreversible brain damage in 1988 as a result of hypoxia from carbon monoxide poisoning (Milner et al., 1991). Studies of DF's visual abilities have shown that DF is unable to report the size, shape, and orientation of an object, either verbally or manually. On the other hand, an analysis of her visuo-motor abilities demonstrates that she shows normal pre-shaping and rotation of her hand when reaching out to grasp the same objects. In other words, DF is able to use visual information about the location, size, shape, and orientation of objects to control her grasping movements (and other visually guided movements), despite the fact that she is unable to perceive and report those same object features. As a concrete example of this dissociation, DF can orient her hand correctly to grasp a rectangular plaque that is placed in front of her at different orientations, as assessed by biomechanical measurements. However, when presented with those same plaques and asked to report the orientation without acting on the object, she cannot do it (Goodale et al., 1991). In fact, she cannot even judge whether a grating stimulus presented on a computer screen is vertical or horizontal (Humphrey et al., 1995).

Structural MRI of DF's brain shows that her lesion is relatively focal and is located bilaterally in an area known as the lateral occipital complex (LOC) (James et al., 2003). The LOC has been studied in healthy individuals for almost 15 years. The LOC is a region on the lateral surface of the cortex at the junction of the occipital and temporal lobes (Malach et al., 1995). The results of many studies have confirmed that the LOC produces more activation with intact objects than with any other class of control stimuli tested. Although the mechanisms of object recognition are unknown and remain the source of intense study, researchers agree that the LOC is intimately involved in visual object recognition, categorization, and naming (Grill-Spector et al., 2001; James et al., 2000; Kourtzi et al., 2003). The location of DF's lesion site and her inability to recognize objects suggests not only that the LOC is involved in visual object recognition but also that an intact LOC is *necessary* for recognition of objects.

Although the lesion to DF's occipito-temporal cortex was relatively focal, there was evidence of smaller scale atrophy throughout the brain, as indicated by enlarged sulci and ventricles (James et al., 2003). Despite the abnormal structural appearance of the atrophied regions, however, BOLD responses from those regions appeared normal. This was in dramatic contrast with the lesion sites, in which the signal resembled those from cerebral spinal fluid, that is, there was no BOLD response. One of the areas that showed striking similarity of functional response between DF and healthy control subjects was the intraparietal sulcus (IPS). In humans and other primates, the IPS has been implicated in planning object-directed actions and spatially representing the environment (Culham and Valyear, 2006; Culham et al., 2006; Frey et al., 2005; Grefkes and Fink, 2005). DF is able to use visual input to guide her object-directed actions, such as reaching and grasping. When she performed grasping actions in the MRI, BOLD activation in her IPS region resembled activation seen in healthy control subjects (James et al., 2003). Thus, DF's case provides clear and converging evidence from behavioral and neuroimaging measures for a separation of function between the dorsal and the ventral visual cortical streams.

The idea of two separate processing pathways, either the what–where hypothesis or the action–perception hypothesis, has been extremely influential in the study of vision, but it has also influenced the study of other sensory systems. A dual-streams approach has been adopted by other researchers to explain the organization of the auditory system (Arnott et al., 2004; Hickok and Poeppel, 2004; Saur et al., 2008) and somatosensory system (Dijkerman and de Haan, 2007; James et al., 2007; Reed et al., 2005).

13.3 Converging Visual and Somatosensory Pathways

Like the visual system, the somatosensory system is organized hierarchically and potentially into two or more separate pathways. There are at least three different two-stream hypotheses that describe the organization of the neural substrates involved in haptic exploration of the environment and, specifically, exploration of

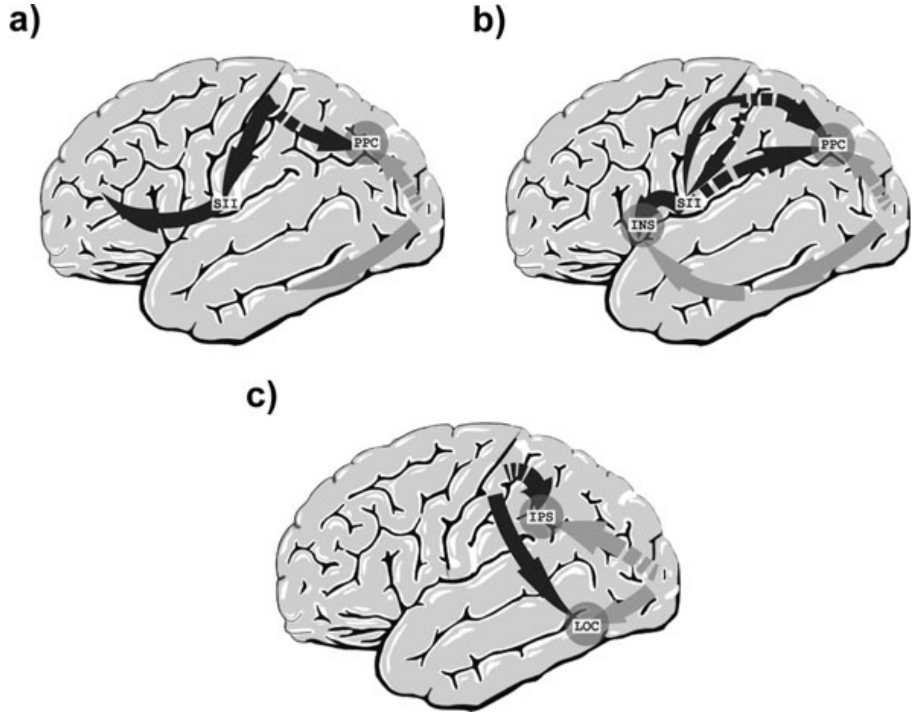

Fig. 13.1 Three two-stream models of haptic object processing. *Light gray arrows* represent visual streams originating in posterior occipital cortex. *Dark gray arrows* represent somatosensory streams originating in post-central cortex. *Solid arrows* denote ventral stream projections, whereas *dashed arrows* denote dorsal stream projections. *Transparent circles* represent zones of multisensory convergence. Abbreviations: LOC, lateral occipital cortex; IPS, intraparietal sulcus; INS, insula; PPC, posterior parietal cortex, SII, secondary somatosensory cortex

physical objects (Dijkerman and de Haan, 2007; James et al., 2007; Reed et al., 2005). The first model (Fig. 13.1a) contends that haptic signals that code object identity are processed in a ventral pathway, which projects from primary somatosensory cortex to the inferior parietal lobe and prefrontal cortex, whereas haptic signals that code the spatial location of objects are processed in a dorsal pathway, which projects from primary somatosensory cortex to the superior parietal lobe (Reed et al., 2005). This theory is similar to the what–where hypothesis in vision, separating the processing of identity from the processing of location (Ungerleider and Mishkin, 1982).

The second model (Fig. 13.1b) contends that processing of haptic signals for object recognition and perception is accomplished by a ventral pathway, which projects from primary and secondary somatosensory cortex to the insula, whereas processing of haptic signals for motor actions is accomplished by a dorsal pathway, which projects from primary and secondary somatosensory cortex to the posterior parietal cortex (Dijkerman and de Haan, 2007). This theory is similar to the visual action–perception hypothesis, separating the processing of objects for perceptual judgments and the processing of objects for actions. An important consideration in this model is the *interaction* between the dorsal and the ventral pathways. Although the visual action–perception hypothesis allows for interactions between the dorsal and the ventral streams, Dijkerman and colleagues suggest that these interactions should be stronger for haptic streams than for vision. A second consideration is the formation of a body representation or body sense. Representations in vision would tend to be of the external environment, even though the receptors are located within the body. Representations of one's own body are completely internal, which may make this type of representation different from visual representations. A third consideration is that the dorsal and ventral streams for vision and haptics may have sites of convergence within the cortex, that is, the pathways may converge at specific sites to integrate information, and these sites may be specific to dorsal and ventral stream function.

The third model (Fig. 13.1c) specifically addresses the convergence of two haptic and two visual streams (James et al., 2007). Haptic object processing is organized into two streams: a ventral stream that projects from the primary somatosensory cortex to the lateral occipito-temporal cortex (LOC) and a dorsal stream that projects from primary somatosensory cortex to the posterior parietal cortex, specifically the intraparietal sulcus (IPS). These two cortical sites (LOC and IPS) mark points of convergence between the visual and the haptic streams of object processing. Convergence of the dorsal visual and haptic pathways is specialized for processing multisensory shape cues to plan object-directed motor actions, whereas convergence of the ventral visual and haptic pathways is specialized for processing multisensory shape cues for object perception, which in turn allows for deliberation and reasoning about our choice of actions directed toward those objects.

Findings to support these models come from a combination of behavioural and neuroimaging studies with patients and healthy subjects, and neurophysiological single-unit recording in nonhuman primates. In the late 1990s, a group of cross-modal haptic–visual priming studies (Easton et al., 1997a, b; Reales and Ballesteros, 1999) changed conceptions in cognitive psychology of how object shape may be represented. In these studies, previous experience with an object facilitated subsequent performance when naming that object. The important finding was that (in most cases) the facilitation occurred whether the sensory modalities of the initial experience and the subsequent test matched or mismatched. The findings of these studies, and subsequent studies (Newell et al., 2001; Norman et al., 2004), suggested that a common representation of shape was used by vision and haptics.

Subsequently, several fMRI studies demonstrated that haptic object recognition recruited areas of putative visual cortex (Amedi et al., 2001; James et al., 2002; Reed et al., 2004; Sathian et al., 1997), suggesting an overlap of visual and haptic *neural* representations for objects. The visual cortical region most consistently recruited in these studies was a sub-region of the LOC, which has been labeled by some as LOtv, for tactile–visual (Amedi et al., 2002). The visual and somatosensory systems process many characteristics of objects. One of the most salient characteristics of objects is their shape (texture is also salient, but shape seems to be the most salient). The results of a number of studies suggest that the key characteristic of objects that leads to recruitment of LOtv is their shape (Amedi et al., 2007; James et al., 2002; Stilla and Sathian, 2008). Some studies also suggest that shape is the most important characteristic for IPS (Culham et al., 2006; Grefkes et al., 2002; Kitada et al., 2006; Peltier et al., 2007). Thus, the processing in LOtv (and perhaps IPS) may not be "visual" or even "visuo-haptic," but instead may be "metamodal." In other words, *shape* information may be processed in LOtv regardless of input modality (Amedi et al., 2007; Pascual-Leone and Hamilton, 2001).

Behavioral data collected from patients with visual agnosia converge with these neuroimaging findings (Feinberg et al., 1986; James et al., 2005; Morin et al., 1984; Ohtake et al., 2001). For instance, patient DF's lesion in the occipito-temporal cortex overlaps with the location of LOC in healthy subjects and has impaired her ability to name or match objects visually, especially when those judgments must be made based on an object's shape (Humphrey et al., 1994; James et al., 2003). However, DF's haptic object recognition ability is also impaired compared to healthy subjects. On three separate tasks, old/new recognition, sequential matching, and paired associates, DF was equally impaired when using vision or haptics (James et al., 2005). In addition to these patient lesion data, transcranial magnetic stimulation (TMS) has been used to produce "transient virtual lesions" in the cortex of healthy individuals. Disrupting neural processing in the occipital cortex caused impairments on a tactile orientation discrimination task (Zangaladze et al., 1999). Taken together, these behavioral and neuroimaging findings from patients and healthy subjects converge to suggest that the visual and haptic systems share overlapping neural substrates for object recognition based on shape analysis (Amedi et al., 2005; James et al., 2007). The most likely candidate for that neural substrate is the LOtv, which resides in what is considered the ventral perceptual stream of visual processing.

Even more intuitive than the convergence of vision and haptics for the recognition of objects is the convergence of vision and haptics for manual interaction with objects. Haptic feedback plays a large role in the calibration of visuo-motor actions (Coats et al., 2008). Neuroimaging studies have shown that at least one area of the intraparietal sulcus (IPS) is involved in bi-modal visuo-haptic processing of shape or geometric properties of objects (Bodegard et al., 2001; Culham and Kanwisher, 2001; Grefkes et al., 2002; Peltier et al., 2007; Roland et al., 1998; Zhang et al., 2004). The IPS is also intimately involved in the planning and execution of sensorimotor actions, including eye movements, pointing, reaching, and grasping (Culham and Valyear, 2006; Grefkes and Fink, 2005). Of particular relevance are the anterior and middle aspects of the IPS. The functional significance of these areas is broad, including the preparation of grasping movements, sensitivity to visual or haptic input, and the processing of object shape and size. Patients with damage to IPS can suffer from tactile apraxia, which is characterized by an inability to recognize objects haptically due to inappropriate use of exploratory movements (Binkofski et al., 1998, 2001; Pause, 1989). These converging lines of evidence have led researchers to conclude that IPS is a site of convergence for several inter-related sensorimotor processes that rely on visual, haptic, and motor information to analyze object shape.

Data from nonhuman primate single-unit recordings strongly support the claim for bi-modal visual and somatosensory processing in IPS (Buneo et al., 2002; Murata et al., 2000; Taira et al., 1990), which is considered homologous with IPS in humans (Grefkes and Fink, 2005; Grefkes et al., 2002). Support for bi-modal processing in LOtv from single-unit recording data, however, is much less sure. One issue is that a nonhuman primate homologue for LOtv has not been established (Tootell et al., 2003). But, despite the issue of homology, there is some evidence that neurons in the ventral visual stream of nonhuman primates do receive both visual and tactile inputs (Maunsell et al., 1991). Thus, the current state of research into visual and haptic shape processing pathways in humans suggests that each sensory system has two functionally specialized cortical pathways, and these pathways converge on at least two separate cortical locations, LOtv and IPS. LOtv is involved in visual and haptic processing of shape information for the purpose of recognition and perception, whereas IPS is responsible for processing visual and haptic shape information for the purposes of guiding object-directed actions. What has not been directly tested, however, is the manner in which signals from the two sensory inputs are combined or integrated in these areas.

13.4 Measuring Neuronal Convergence with BOLD fMRI

When describing the convergence of sensory inputs onto brain regions, researchers in the field of multisensory neurophysiology distinguish between two types of convergence: areal convergence and neuronal convergence (Meredith, 2002). In describing the research on convergence of visual and haptic inputs in the preceding sections, this distinction was not made. *Areal convergence* describes the case when different sensory inputs project to neurons in the same brain region, but do not synapse on the exact same neurons. Because the inputs do not synapse on the same neurons, there is no interaction or integration of the inputs. *Neuronal convergence*, on the other hand, describes the case when inputs from different sensory systems project to the same neurons. By synapsing on the same neurons, the inputs interact at the neural level and can be integrated. Areal convergence and neuronal convergence are relatively simple to dissociate with single-unit recording. If a neuron changes its activity when the animal is simultaneously stimulated through two sensory modalities compared to only one, then the neuron is integrating those inputs (Meredith and Stein, 1983; Stein and Stanford, 2008). Because single-unit recording is difficult or impossible to perform in humans, multisensory integration in the human brain has been investigated using functional neuroimaging techniques. Because BOLD fMRI activation is measured from clusters of voxels that represent large populations of neurons, distinguishing between areal and neuronal convergence with fMRI invokes a different set of criteria than with single units. Because fMRI is a newer methodology than single-unit recording, these criteria are not as well established.

One issue with predicting the strength BOLD activation with multisensory stimuli is that populations of neurons in known multisensory cortical regions contain unisensory as well as multisensory neurons (Barraclough et al., 2005; Benevento et al., 1977; Hikosaka et al., 1988). Under this assumption, the null hypothesis to be rejected is that a multisensory stimulus produces activation equivalent to the sum of that produced by the unisensory components (Calvert et al., 2000; Laurienti et al., 2005). This is because the combination stimulus should excite the unisensory neurons *at least* as effectively as the component stimuli. Only if the combination stimulus produces more activation than this additive null hypothesis ("superadditivity"), do the results imply an interaction between sensory streams. Based on the known neurophysiology, the most likely interpretation of an interaction is that there is a third pool of multisensory neurons in the population, in addition to the two unisensory pools. To be clear, this hypothetical third pool of neurons includes any neuron with a response that is not considered unisensory. Thus, the third pool includes all types of multisensory neurons, including those that show multisensory enhancement and those that show suppression, those with linear or additive responses with multisensory stimuli, and those with nonlinear responses with multisensory stimuli. Known sites of multisensory convergence, such as the superior temporal sulcus (STS) for audio-visual stimuli, have many different types of multisensory neurons combined with unisensory neurons. Thus, a site like STS should produce a pattern of activation that rejects the null hypothesis of only two pools of unisensory neurons. In practice, though, known sites of multisensory convergence like STS rarely show statistically significant evidence of superadditivity with BOLD fMRI signals (Beauchamp, 2005b; Stevenson et al., 2007). Thus, other factors must play a role in determining whether or not BOLD fMRI measurements can detect the presence of a pool of multisensory neurons.

It is generally understood that BOLD fMRI measurements lack a natural zero value or well-defined baseline (Raichle et al., 2001; Stark and Squire, 2001). Because of this constraint, BOLD is considered a "relative" measure of neural activation, rather than an absolute measure. In other words, only *differences* in BOLD response between conditions are meaningful, not the absolute levels. It is possible that the use of absolute BOLD values in the calculation of an additive criterion has led to a lack of consistency across studies assessing multisensory integration of specific brain regions.

The influence of an arbitrary baseline on the superadditive criterion is illustrated graphically in Fig. 13.2. The two top graphs show raw BOLD data collected in two different experiments. Because this is a simulation, we can make the data in the two experiments the same, except for one important factor, which is that experimenters for the different experiments have chosen to use different baselines (or have done so unknowingly). In Experiment 1, the baseline is slightly below the "true baseline" (natural zero), and in Experiment 2, the baseline is slightly above the "true baseline." Raw BOLD activation for the multisensory stimulus condition (VH) is simulated based on a neural population composed of *only unisensory visual and unisensory haptic neurons*. Thus, the result of both experiments should be to *fail to reject* the null hypothesis. The established practice in fMRI is to convert raw BOLD values into percent signal change values by subtracting the value of the baseline condition and then dividing by it. The bottom two graphs in Fig. 13.2 show the data from the top two graphs after undergoing this transformation. Recall that the only difference between experiments was the different baseline activation; therefore, differences between the left and right bottom graphs are due only to a difference in the

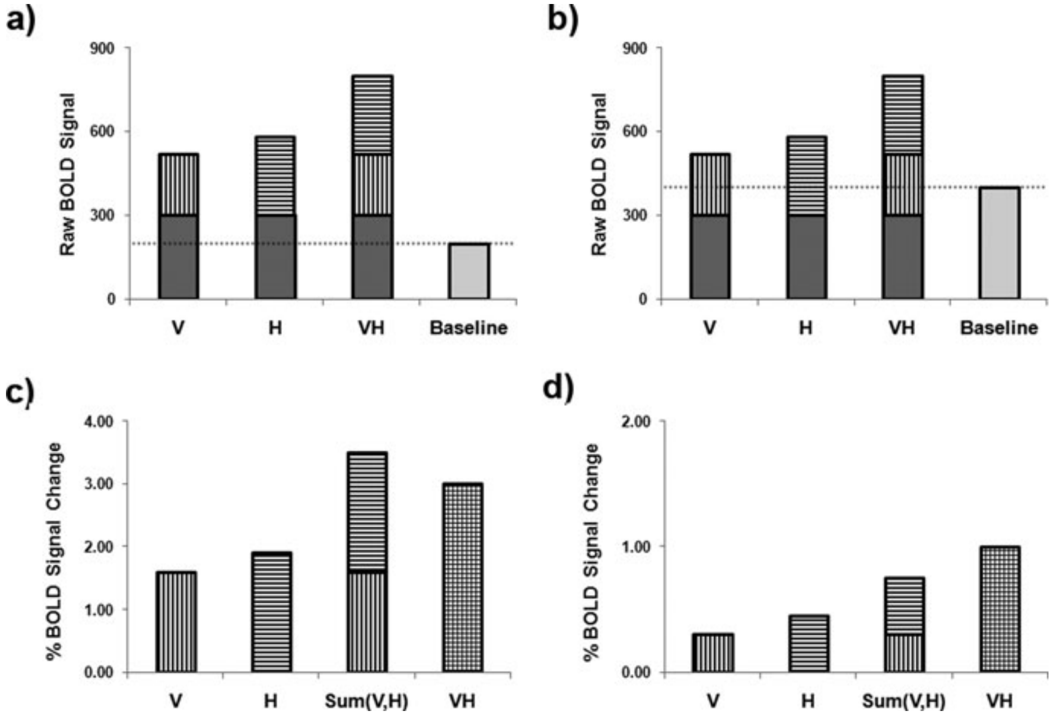

Fig. 13.2 The influence of baseline activation on the absolute additive criterion. (**a**) and (**b**) The height of stacked bars indicates the contribution of different factors to the raw BOLD signal. *Dark gray bars* indicate an arbitrary value added during the reconstruction of MRI images. *Horizontal* and *vertical lines* indicate the contribution of neural activity from visual (V) or haptic (H) sensory channels, respectively. The *light gray bar* is the BOLD activation produced by the "baseline" condition. (**c**) and (**d**) Percent BOLD signal change calculated based on the different baseline activation values shown in (**a**) and (**b**), respectively. Signal change values are proportional to the difference between the total height of the stacked bars and the dotted line indicating the level of baseline activation. The Sum (V,H) is the absolute additive criterion. (**c**) Superadditivity and (**d**) Subadditivity

baseline. The signal change values in the two graphs are clearly different. Both of these experiments would reject the null hypothesis, but the effects are in the opposite direction. More alarming is that rejecting the null hypothesis would reveal nothing about the underlying neural populations, but is completely dependent on the activation of the baseline condition. It is possible that this type of influence may explain the inconsistency in results from different research groups using superadditivity as a criterion (Beauchamp, 2005a, b; Beauchamp et al., 2004a, b; Laurienti et al., 2006; Peiffer et al., 2007; Stevenson and James, 2009; Stevenson et al., 2007).

Because absolute BOLD measurements produce inconsistent results when used as a criterion for the assessment of multisensory integration, we have recently turned to using relative *differences* in BOLD activation. The use of relative differences alleviates the problem of an indeterminate baseline, because the baseline components embedded in the two measurements cancel out when a difference operation is performed. The null hypothesis for these BOLD differences is similar to absolute BOLD measurements and follows a similar hypothesis of additivity, that is, the null hypothesis is the sum of the two unisensory *differences*. If the multisensory *difference* differs from the null, one can infer an interaction between sensory channels in the form of a third pool of multisensory neurons using the same logic applied to the superadditive null hypothesis. The benefit of using differences, however, is that they are not susceptible to changes in baseline.

BOLD differences can be calculated across any manipulation of the stimulus or task that produces a systematic, monotonic change in BOLD activation. For instance, BOLD differences have been successfully used with audio-visual stimuli for which the signal-to-noise ratio (SNR) was varied (Stevenson and James, 2009). In that study, unisensory audio and visual stimuli produced less BOLD activation in the superior temporal sulcus (STS) as stimuli were degraded by lowering the SNR. Multisensory stimuli also showed less activation with reduced SNR, but the decrease in activation was not as large as predicted by the null hypothesis. That is, the BOLD difference for multisensory stimuli was less than the sum of the two unisensory differences. This effect on multisensory BOLD activation was called inverse effectiveness, because it resembled an effect often seen in single-unit recordings taken from multisensory regions. As stimuli are degraded, they are less effective at stimulating unisensory and multisensory neurons. Because the relative multisensory gain increases as effectiveness decreases, the effect is called *inverse effectiveness* (Meredith and Stein, 1986). Although a pattern of inverse effectiveness in BOLD activation does not necessarily imply that neurons in that area are inversely effective, it does imply a difference from the null hypothesis, and thus an interaction between sensory channels.

13.5 Sites of Visuo-haptic Neuronal Convergence

Although there is considerable evidence for bi-modal visual and haptic processing of object shape in the primate brain in at least two cortical sites, LOtv and IPS, until very recently (Tal and Amedi, 2009) a test for neuronal convergence in humans using fMRI had not been reported. Using the BOLD differences method described above, we designed an experiment to test for neuronal convergence of visual and haptic inputs in the LOC and IPS (Kim and James in press). Our stimulus manipulation was to vary the level of stimulus quality. Finding a significant “difference of differences” across levels of stimulus quality would confirm the presence of multisensory integration, even in the absence of superadditivity. Based on previous results with audio-visual integration using this method, we hypothesized that the direction of that difference would be in the direction predicted by inverse effectiveness, that is, as stimulus quality was reduced, there should be a smaller drop in multisensory activation than predicted based on the drop in unisensory activation.

We localized ROIs using a standard method that contrasts visual objects (VO) with visual textures (VT) and haptic objects (HO) with haptic textures (HT) (Amedi et al., 2001). The visual contrast is the same standard functional localizer used to isolate the LOC visually (Malach et al., 1995). The haptic contrast typically activates a large cluster along the entire IPS. A conjunction of the two contrasts isolates overlapping regions that are object selective for both sensory modalities. In the ventral stream, the overlapping region is labeled LOtv and is found consistently with the conjunction contrast. An overlapping region in the dorsal stream is less consistently found with the conjunction contrast (Amedi et al., 2005; Lacey et al., 2009), perhaps because researchers do not control the parameters of active exploration. Nevertheless, our use of the conjunction contrast did find statistically reliable clusters of voxels in the IPS. Figure 13.3 shows the two functional ROIs, LOtv and IPS, localized on the group-average functional data ($N = 7$) and superimposed on group-average anatomical images. The analysis used to produce the maps was a conjunction of four contrasts, HO–HT and VO–VT and HO–VT and VO–HT. This analysis localizes regions that are bi-modal and object selective and that also have equal contribution from visual and haptic conditions. Images on the left and right of Fig. 13.3 are shown at different statistical thresholds. At the more conservative threshold (Fig. 13.3a, c), the right-hemisphere cluster does not survive. At the more liberal threshold (Fig. 13.3b, d), the left-hemisphere cluster is much larger than the right. The same left-hemisphere bias for haptic or visuo-haptic object processing has been shown in previous studies; however, the significance of the possible lateralization of functional is unknown. We wanted to analyze both the left and the right hemispheres and wanted to equate the size (i.e., number of voxels) of the left- and right-hemisphere clusters. Due to the reliable left-hemisphere bias, this

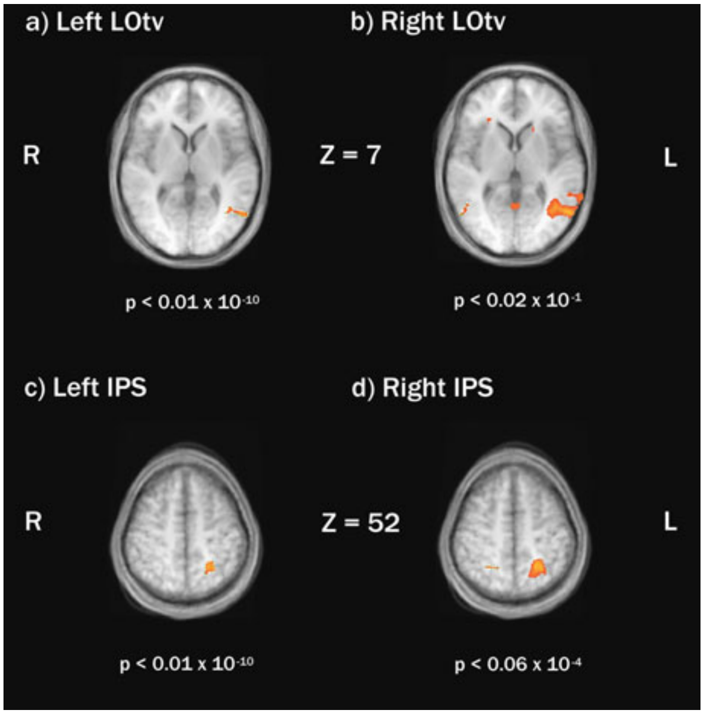

Fig. 13.3 Bi-modal visuo-haptic object-selective regions of interest. The boundaries of the regions of interest used in the analyses are outlined in *bright yellow*. The heights of the axial slices in Talairach space are indicated by the Z = labels

meant using a different statistical threshold for determining the left- and right-hemisphere ROIs. Thus, the left-hemisphere ROIs were the yellow-outlined clusters in Panels A and C and the right-hemisphere ROIs were the yellow-outlined clusters in Panels B and D. The large, non-outlined clusters in the left hemisphere in Panels B and D were not analyzed. Importantly, the threshold for the left-hemisphere ROIs was set at a typically conservative value. A more liberal threshold was used only for the right-hemisphere ROI. The results of these contrasts were consistent with previous research suggesting that LOtv and IPS are instrumental in processing object shape information (Amedi et al., 2005). They are also consistent with previous work suggesting that LOtv and IPS are sites of convergence for visual and haptic sensory inputs (James et al., 2007).

To assess inverse effectiveness and superadditivity in these ROIs, we presented subjects with novel objects from two categories and instructed them to perform a two-alternative forced-choice decision. Sixteen objects were used. All objects were created by attaching four wooden geon-like geometric components in a standard configuration to provide differences in shape information (Biederman, 1987). In this experiment, only one of the components was diagnostic of category membership. Eight objects in Category 1 had a half-circle-shaped diagnostic component, and the other eight objects in Category 2 had a triangle-shaped diagnostic component (Fig. 13.4a). Each stimulus was 14 cm wide and 9.5 cm long. Texture on the stimuli was determined by the size of nylon beads that were glued to the surface. Textures and non-diagnostic features could not be used to perform the 2AFC task. The purpose of the different textures and non-diagnostic shape components was to add complexity and variability to the psychological object similarity space and keep subjects more interested in the task, but subjects were explicitly instructed to use only the diagnostic shape feature to perform the task. The distribution of textures and non-diagnostic components was the same across the two object categories.

For visual presentation, a grayscale picture of each stimulus was presented using an LCD projector and rear-projection screen. Subjects viewed the images using a rear-face mirror attached to the head coil. Visual stimuli were presented at 12° × 8°

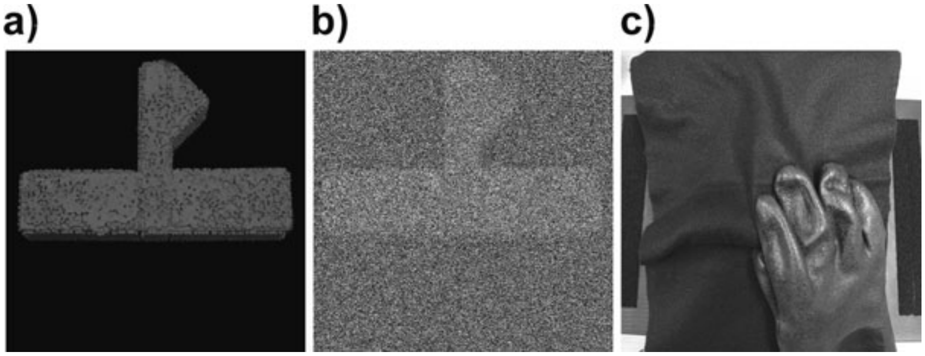

Fig. 13.4 Examples of procedures for degrading stimuli in visual and haptic conditions. Pictures of an undegraded object (**a**), a visually degraded object (**b**), and a haptically degraded object being palpated (**c**)

of visual angle. To establish different levels of salience, a fixed level of external Gaussian noise was added to the stimuli (Fig. 13.4b) and the stimulus signal contrast was reduced to two different levels. These levels were calibrated independently for each individual and reflected their 71 and 89% performance thresholds.

For haptic presentation, an angled "table" was placed on the subject's midsection. An experimenter standing in the MRI room delivered tangible stimuli to the subject by placing them on a designated spot on the table. Subjects palpated the objects with eyes closed with both hands and were instructed not to lift the objects from the table. To establish different levels of salience, subjects wore a pair of PVC gloves, which reduced their tactile sensitivity. Individual 71 and 89% performance thresholds were measured by covering objects with a different number of layers of thick felt fabric (Fig. 13.4c). The layered fabric further reduced tactile sensitivity, but unlike the gloves, allowed the experimenter to rapidly change between performance (or sensitivity) levels.

Subjects and the experimenter both wore headphones and listened for a sequence of auditory cues that indicated when subjects should start and stop hand movements and when the experimenter should switch out the stimulus. The final design had two factors: stimulus quality and stimulus modality, with two levels of quality (high and low) and three modalities (visual [V], haptic [H], and visuo-haptic [VH]). Seven subjects participated in the experiment. Imaging parameters and data pre-processing steps were standard and are described elsewhere (Kim and James in press). Accuracy and reaction time measures showed strong effects of stimulus quality. In the unisensory conditions, accuracy levels were close to 71% for low quality and 89% for high quality, the performance levels for which the conditions were calibrated. Reaction times were also slower for the low quality than high quality condition.

BOLD activations in left and right LOtv and IPS are shown in Fig. 13.5 for all conditions in the 2×3 design. The null hypothesis for the additive model is also presented for each level of stimulus quality and labeled S(V,H). It is apparent from comparing the VH stimulus condition to the null hypothesis that there is no evidence of superadditivity in either LOtv or IPS. The VH stimulus condition produced subadditive activation in all cases except for the high quality condition in right LOtv, which was additive. Thus, if superadditivity were the only criterion, we would infer from these data that LOtv and IPS do not represent sites of neuronal convergence for visual and haptic sensory inputs.

Figure 13.6 shows the results of the new BOLD differences analysis, performed on the same data shown in Fig. 13.5. Instead of comparing absolute levels of BOLD activation, recall that this analysis compares *differences* in BOLD activation. Thus, the height of the bars in Fig. 13.6 represents the difference in BOLD activation between high- and low-signal quality conditions. This difference is represented for visual (ΔV), haptic (ΔH), and visuo-haptic (ΔVH) stimulus conditions. The null hypothesis to be rejected is represented by the sum of unisensory differences (i.e., S(ΔV, ΔH) bar). There is a clear difference between the ΔVH and S(ΔV, ΔH) for both LOtv and IPS in the left hemisphere. The difference in the right hemisphere is in the same direction, but is less robust. Based on this rejection of the

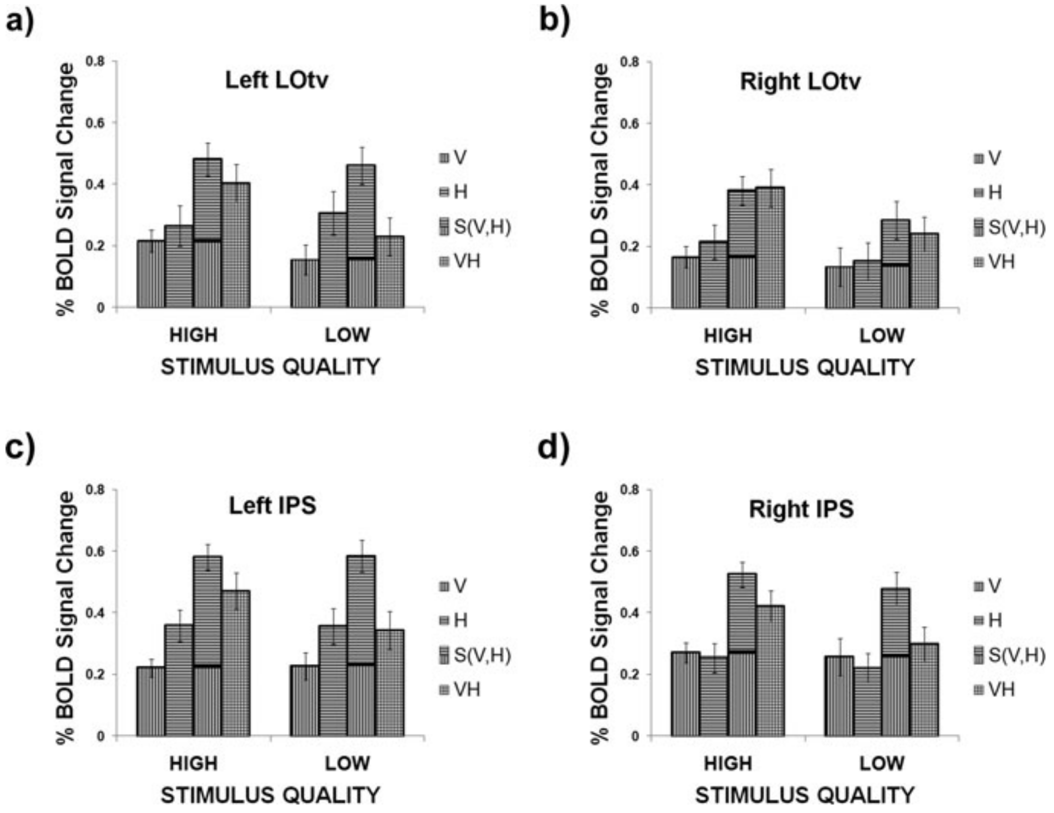

Fig. 13.5 BOLD percent signal change as a function of sensory condition, stimulus quality, region of interest, and hemisphere shown for the following regions: a) left LOtv, b) right LOtv, c) left IPS and d) right IPS. S(V, H) represented the absolute additive criterion

null hypothesis, we can infer that the underlying neuronal population is not composed of two pools of unisensory neurons: one visual and one haptic. Based on the neurophysiology of known multisensory areas, a likely inference is that these two areas contain a mixture of unisensory and multisensory neurons (Meredith and Stein, 1983, 1986).

Rejection of our null hypothesis, though, simply means that there was a difference between ΔVH and the sum of ΔV and ΔH. A difference in either direction implies an interaction between sensory modalities, possibly due to a third pool of multisensory neurons. But, the direction shown in Fig. 13.6 was unexpected. One of the general principles of single-unit recordings from multisensory neurons is that multisensory enhancement increases as the stimuli are degraded in quality. That is, the gain in activity with a multisensory stimulus over and above the activity with a unisensory stimulus increases with decreasing quality. Low-quality stimuli are also less effective at stimulating unisensory and multisensory neurons. Thus, inverse effectiveness describes the phenomenon that as the effectiveness of a stimulus decreases, the multisensory gain increases. Given that inverse effectiveness is a known principle of neural activity in multisensory neurons, we predicted that activity in LOtv and IPS, if the null hypothesis was rejected, would also show evidence of inverse effectiveness. However, the pattern of change shown in Figs. 13.5 and 13.6

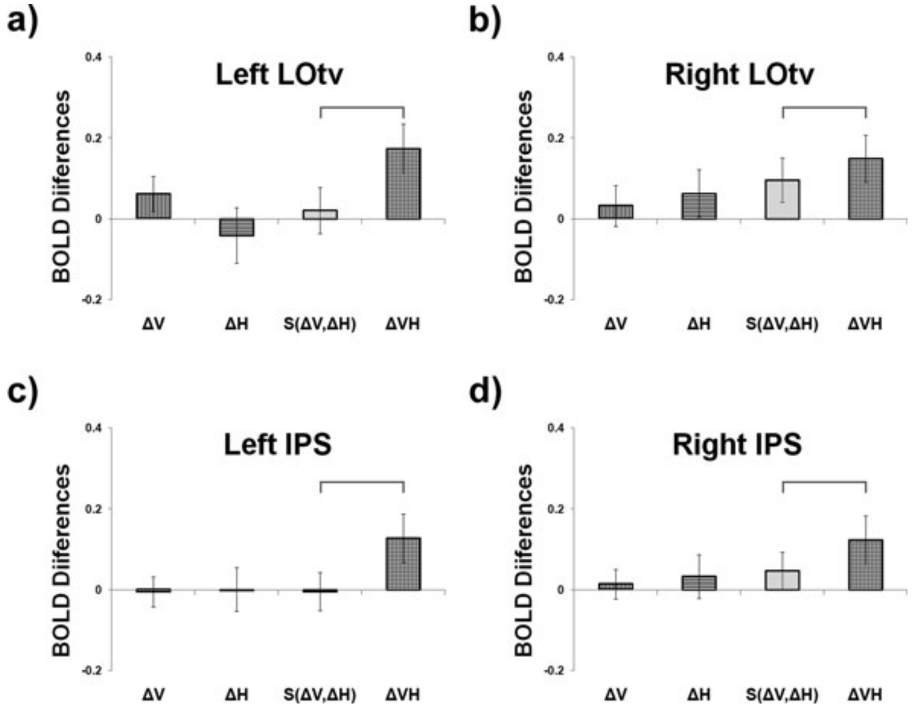

Fig. 13.6 BOLD differences as a function of sensory condition, region of interest, and hemisphere shown for the following regions: a) left LOtv, b) right LOtv, c) left IPS and d) right IPS. Differences were calculated as high quality minus low quality. S(ΔV,ΔH) represents the additive-differences criterion. ΔVH represents the multisensory difference, which is compared to S(ΔV,ΔH) to establish the presence of enhanced or inverse effectiveness

is the opposite. The multisensory gain is stronger for high-quality stimuli than for low-quality stimuli. It is important to note that the opposite direction of the effect is not due to an indirect relation between stimulus quality and brain activation (effectiveness). That is, if high-quality stimuli produced less activation in these regions than low-quality stimuli, then that alone could make the change in gain appear to go in the opposite direction.

The multisensory gains shown in Figs. 13.5 and 13.6 are clearly stronger with the high-quality stimuli. This effect is the opposite of that seen with inverse effectiveness (Meredith and Stein, 1986). Thus, we suggest that this effect should be called "enhanced effectiveness" (Kim and James in press), because as the effectiveness of the stimuli is enhanced, the multisensory gain is also disproportionately enhanced.

Both LOtv and IPS showed evidence for integration of visual and haptic sensory inputs and both LOtv and IPS showed enhanced effectiveness. The effect was in the same direction in both the left and the right hemispheres, but was much stronger in the left. The results provide further evidence that the visual and haptic systems process object shape through two pathways and that LOtv and IPS represent points of convergence for those two pathways (James et al., 2007). LOtv forms part of the ventral or perceptual pathway for vision and haptics, and IPS forms part of the dorsal or action pathway for vision and haptics. The results suggest that integration of visual and haptic sensory inputs is *similar* in LOtv and IPS. This result may be unexpected, given the emphasis placed on separable processes in the two pathways (Goodale et al., 1991). However, our results suggest only that the properties of multisensory convergence are similar in the two regions, not that the underlying neural processes that rely on that convergence are the same. Although speculative, our data may suggest that the properties of multisensory convergence are similar in different areas of cortex. To more directly test this hypothesis, future studies of neuronal convergence of visual and haptic sensory channels in ventral and dorsal pathways should investigate both object recognition and object-directed actions.

13.6 Conclusions

We have reviewed the evidence for separable action and perception pathways in both the visual and the haptic systems for the analysis of object shape. These systems converge on at least two neural sites: one in the dorsal action pathway and one in the ventral perception pathway. We tested these sites of convergence for evidence of multisensory integration of visual and haptic inputs. In the absence of superadditivity, we found evidence for multisensory integration (neuronal convergence) using a new method that employs relative BOLD differences instead of absolute BOLD values. Dorsal and ventral sites showed the same general pattern of multisensory integration. This suggests that integration of object shape information in the dorsal and ventral streams may occur by the same general mechanisms.

Acknowledgments This research was supported in part by the Faculty Research Support Program, administered by the Indiana University Office of the Vice President of Research, and in part by the Indiana METACyt Initiative of Indiana University, funded in part through a major grant from the Lilly Endowment, Inc.

References

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, 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, Stern WM, Camprodon JA, Bermpohl F, Merabet L, Rotman S, Hemond C, Meijer P, Pascual-Leone A (2007) Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. Nat Neurosci 10:687–689

Amedi A, von Kriegstein K, van Atteveldt NM, Beauchamp MS, Naumer MJ (2005) Functional imaging of human crossmodal indentification and object recognition. Exp Brain Res 166: 559–571

Arnott SR, Binns MA, Grady CL, Alain C (2004) Assessing the auditory dual-pathway model in humans. Neuroimage 22:401–408

Baizer JS, Ungerleider LG, Desimone R (1991) Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. J Neurosci 11:168–190

Barraclough NE, Xiao D, Baker CI, Oram MW, Perrett DI (2005) Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J Cogn Neurosci 17:377–391

Beauchamp MS (2005a) See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. Curr Opin Neurobiol 15:145–153

Beauchamp MS (2005b) Statistical criteria in FMRI studies of multisensory integration. Neuroinformatics 3:93–113

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004a) Unraveling multisensory integration: patchy organization within human STS multisensory cortex. Nat Neurosci 7: 1190–1192

Beauchamp MS, Lee KE, Argall BD, Martin A (2004b) Integration of auditory and visual information about objects in superior temporal sulcus. Neuron 41:809–823

Benevento LA, Fallon J, Davis BJ, Rezak M (1977) Auditory--visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. Exp Neurol 57:849–872

Biederman I (1987) Recognition-by-components: a theory of human image understanding. Psychol Rev 94:115–147

Binkofski F, Dohle C, Posse S, Stephan KM, Hefter H, Seitz RJ, Freund HJ (1998) Human anterior intraparietal area subserves prehension: a combined lesion and functional MRI activation study. Neurology 50:1253–1259

Binkofski F, Kunesch E, Classen J, Seitz RJ, Freund HJ (2001) Tactile apraxia: unimodal apractic disorder of tactile object exploration associated with parietal lobe lesions. Brain 124: 132–144

Bodegard A, Geyer S, Grefkes C, Zilles K, Roland PE (2001) Hierarchical processing of tactile shape in the human brain. Neuron 31:317–328

Buneo CA, Jarvis MR, Batista AP, Andersen RA (2002) Direct visuomotor transformations for reaching. Nature 416:632–636

Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr Biol 10:649–657

Clark A (1997) Being there. MIT Press, Cambridge

Coats R, Bingham GP, Mon-Williams M (2008) Calibrating grasp size and reach distance: interactions reveal integral organization of reaching-to-grasp movements. Exp Brain Res 189:211–220

Culham JC, Cavina-Pratesi C, Singhal A (2006) The role of parietal cortex in visuomotor control: what have we learned from neuroimaging? Neuropsychologia 44:2668–2684

Culham JC, Kanwisher NG (2001) Neuroimaging of cognitive functions in human parietal cortex. Curr Opin Neurobiol 11:157–163

Culham JC, Valyear KF (2006) Human parietal cortex in action. Curr Opin Neurobiol 16:205–212

de Sa VR, Ballard DH (1998) Perceptual learning from cross-modal feedback. In: Goldstone RL, Schyns PG, Medin DL (eds) Psychology of learning and motivation, vol 36. Adademic Press, San Diego, CA

Dijkerman HC, de Haan EH (2007) Somatosensory processes subserving perception and action. Behav Brain Sci 30:189–201; discussion 201–139

Easton RD, Greene AJ, Srinivas K (1997a) Transfer between vision and haptics: memory for 2-D patterns and 3-D objects. Psychonomic Bulletin and Review 4:403–410

Easton RD, Srinivas K, Greene AJ (1997b) Do vision and haptics share common representations? Implicit and explicit memory within and between modalities. J Exp Psychol: Learn, Mem, Cogn 23:153–163

Feinberg TE, Gonzalez Rothi LJ, Heilman KM (1986) Multimodal agnosia after unilateral left hemisphere lesion. Neurology 36:864–867

Frey SH, Vinton D, Norlund R, Grafton ST (2005) Cortical topography of human anterior intraparietal cortex active during visually guided grasping. Brain Res Cogn Brain Res 23:397–405

Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends Neurosci 15:20–25

Goodale MA, Milner AD, Jakobson LS, Carey DP (1991) A neurological dissociation between perceiving objects and grasping them. Nature 349:154–156

Grefkes C, Fink GR (2005) The functional organization of the intraparietal sulcus in humans and monkeys. J Anat 207:3–17

Grefkes C, Weiss PH, Zilles K, Fink GR (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, Kourtzi Z, Kanwisher N (2001) The lateral occipital complex and its role in object recognition. Vision Res 41:1409–1422

Hickok G, Poeppel D (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92:67–99

Hikosaka K, Iwai E, Saito H, Tanaka K (1988) Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. J Neurophysiol 60: 1615–1637

Humphrey GK, Goodale MA, Corbetta M, Aglioti S (1995) The McCollough effect reveals orientation discrimination in a case of cortical blindness. Curr Biol 5:545–551

Humphrey GK, Goodale MA, Jakobson LS, Servos P (1994) The role of surface information in object recognition: studies of a visual form agnosic and normal subjects. Perception 23: 1457–1481

James TW, Culham JC, Humphrey GK, Milner AD, Goodale MA (2003) Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. Brain 126: 2463–2475

James TW, Humphrey GK, Gati JS, Menon RS, Goodale MA (2000) The effects of visual object priming on brain activation before and after recognition. Curr Biol 10:1017–1024

James TW, Humphrey GK, Gati JS, Servos P, Menon RS, Goodale MA (2002) Haptic study of three-dimensional objects activates extrastriate visual areas. Neuropsychologia 40:1706–1714

James TW, James KH, Humphrey GK, Goodale MA (2005) Do visual and tactile object representations share the same neural substrate? In: Heller MA, Ballesteros S (eds) Touch and blindness: psychology and neuroscience. Lawrence Erlbaum, Mahwah, NJ

James TW, Kim S, Fisher JS (2007) The neural basis of haptic object processing. Can J Exp Psychol 61:219–229

Kim S, James TW (in press) Enhanced effectiveness in visuo-haptic object-selective brain regions with increasing stimulus salience. Hum Brain Mapp

Kitada R, Kito T, Saito DN, Kochiyama T, Matsumura M, Sadato N, Lederman SJ (2006) Multisensory activation of the intraparietal area when classifying grating orientation: a functional magnetic resonance imaging study. J Neurosci 26:7491–7501

Kourtzi Z, Tolias AS, Altmann CF, Augath M, Logothetis NK (2003) Integration of local features into global shapes: monkey and human fMRI studies. Neuron 37:333–346

Lacey S, Tal N, Amedi A, Sathian K (2009) A putative model of multisensory object representation. Brain Topogr 21:269–274

Laurienti PJ, Burdette JH, Maldjian JA, Wallace MT (2006) Enhanced multisensory integration in older adults. Neurobiol Aging 27:1155–1163

Laurienti PJ, Perrault TJ, Stanford TR, Wallace MT, Stein BE (2005) On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. Exp Brain Res 166:289–297

Lennie P (1998) Single units and visual cortical organization. Perception 27:889–935

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci U S A 92:8135–8139

Maunsell JH, Sclar G, Nealey TA, DePriest DD (1991) Extraretinal representations in area V4 in the macaque monkey. Vis Neurosci 7:561–573

Meredith MA (2002) On the neuronal basis for multisensory convergence: a brief overview. Brain Res Cogn Brain Res 14:31–40

Meredith MA, Stein BE (1983) Interactions among converging sensory inputs in the superior colliculus. Science 221:389–391

Meredith MA, Stein BE (1986) Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. J Neurophysiol 56:640–662

Milner AD, Goodale MA (1995) The visual brain in action. Oxford University Press, Oxford, UK

Milner AD, Perrett DI, Johnston RS, Benson RS, Jordan PJ, Heeley TR, Bettucci DW, Mortara D, Mutani F, Terazzi R, Davidson DLW (1991) Perception and action in visual form agnosia. Brain 114:405–428

Molyneux W (1688) Letter to John Locke. In: de Beer ES (ed) The correspondence of John Locke, vol 3. Clarendon Press, Oxford

Morel A, Bullier J (1990) Anatomical segregation of two cortical visual pathways in the macaque monkey. Vis Neurosci 4:555–578

Morin P, Rivrain Y, Eustache F, Lambert J, Courtheoux P (1984) Visual and tactile agnosia. Rev Neurol 140:271–277

Murata A, Gallese V, Luppino G, Kaseda M, Sakata H (2000) Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. J Neurophysiol 83:2580–2601

Newell FN, Ernst MO, Tjan BS, Bulthoff HH (2001) Viewpoint dependence in visual and haptic object recognition. Psychol Sci 12:37–42

Norman JF, Norman HF, Clayton AM, Lianekhammy J, Zielke G (2004) The visual and haptic perception of natural object shape. Percept Psychophys 66:342–351

Ohtake H, Fujii T, Yamadori A, Fujimori M, Hayakawa Y, Suzuki K (2001) The influence of misnaming on object recognition: a case of multimodal agnosia. Cortex 37:175–186

Pascual-Leone A, Hamilton R (2001) The metamodal organization of the brain. Prog Brain Res 134:427–445

Pause M, Kunesch E, Binkofsku F, Freund HJ (1989) Sensorimotor disturbances in patients with lesions of the parietal corex. Brain 112:1599–1625

Peiffer AM, Mozolic JL, Hugenschmidt CE, Laurienti PJ (2007) Age-related multisensory enhancement in a simple audiovisual detection task. Neuroreport 18:1077–1081

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

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. Proc Natl Acad Sci U S A 98:676–682

Reales JM, Ballesteros S (1999) Implicit and explicit memory for visual and haptic objects: corss-modal priming depends on structural descriptions. J Exp Psychol: Learn, Mem, Cogn 25: 644–663

Reed CL, Klatzky RL, Halgren E (2005) What vs. where in touch: an fMRI study. Neuroimage 25:718–726

Reed CL, Shoham S, Halgren E (2004) Neural substrates of tactile object recognition: an fMRI study. Hum Brain Mapp 21:236–246

Roland PE, O′Sullivan B, Kawashima R (1998) Shape and roughness activate different somatosensory areas in the human brain. Proc Natl Acad Sci 95:3295–3300

Sathian K, Zangaladze A, Hoffman JM, Grafton ST (1997) Feeling with the mind's eye. Neuroreport 8:3877–3881

Saur D, Kreher BW, Schnell S, Kummerer D, Kellmeyer P, Vry MS, Umarova R, Musso M, Glauche V, Abel S, Huber W, Rijntjes M, Hennig J, Weiller C (2008) Ventral and dorsal pathways for language. Proc Natl Acad Sci U S A 105:18035–18040

Stark CE, Squire LR (2001) When zero is not zero: the problem of ambiguous baseline conditions in fMRI. Proc Natl Acad Sci U S A 98:12760–12766

Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. Nat Rev Neurosci 9:255–266

Stevenson RA, Geoghegan ML, James TW (2007) Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. Exp Brain Res 179:85–95

Stevenson RA, James TW (2009) Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. Neuroimage 44:1210–1223

Stilla R, Sathian K (2008) Selective visuo-haptic processing of shape and texture. Hum Brain Mapp 29:1123–1138

Taira M, Mine S, Georgopoulos AP, Murata A, Sakata H (1990) Parietal cortex neurons of the monkey related to the visual guidance of hand movement. Exp Brain Res 83:29–36

Tal N, Amedi A (2009) Multisensory visual-tactile object related network in humans: insights gained using a novel crossmodal adaptation approach. Exp Brain Res 198:165–182

Tootell RB, Tsao D, Vanduffel W (2003) Neuroimaging weighs in: Humans meet macaques in "primate" visual cortex. J Neurosci 23:3981–3989

Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield RJ (eds) The analysis of visual behavior. MIT Press, Cambridge, MA, pp 549–586

Young MP (1992) Objective analysis of the topological organization of the primate cortical visual system. Nature 358:152–155

Zangaladze A, Epstein CM, Grafton ST, Sathian K (1999) Involvement of visual cortex in tactile discrimination of orientation. Nature 401:587–590

Zhang M, Weisser VD, Stilla R, Prather SC, Sathian K (2004) Multisensory cortical processing of object shape and its relation to mental imagery. Cogn Affect Behav Neurosci 4:251–259