



# Visuo-haptic multisensory object recognition, categorization, and representation

Simon Lacey<sup>1\*</sup> and K. Sathian<sup>1,2,3,4</sup>

<sup>1</sup> Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA

<sup>2</sup> Department of Rehabilitation Medicine, Emory University School of Medicine, Atlanta, GA, USA

<sup>3</sup> Department of Psychology, Emory University School of Medicine, Atlanta, GA, USA

<sup>4</sup> Rehabilitation Research and Development Center of Excellence, Atlanta Veterans Affairs Medical Center, Decatur, GA, USA

## Edited by:

Chris Fields, New Mexico State University, USA (retired)

## Reviewed by:

Carl M. Gaspar, Hangzhou Normal University, China

Mounia Ziat, Northern Michigan University, USA

## \*Correspondence:

Simon Lacey, Department of Neurology, Emory University School of Medicine, WMB-6000, 101 Woodruff Circle, Atlanta, GA 30322, USA

e-mail: slacey@emory.edu

Visual and haptic unisensory object processing show many similarities in terms of categorization, recognition, and representation. In this review, we discuss how these similarities contribute to multisensory object processing. In particular, we show that similar unisensory visual and haptic representations lead to a shared multisensory representation underlying both cross-modal object recognition and view-independence. This shared representation suggests a common neural substrate and we review several candidate brain regions, previously thought to be specialized for aspects of visual processing, that are now known also to be involved in analogous haptic tasks. Finally, we lay out the evidence for a model of multisensory object recognition in which top-down and bottom-up pathways to the object-selective lateral occipital complex are modulated by object familiarity and individual differences in object and spatial imagery.

**Keywords:** cross-modal, effective connectivity, fMRI, viewpoint dependence, face processing, visual imagery

## INTRODUCTION

Despite the fact that object perception and recognition are invariably multisensory processes in real life, the haptic modality was for a long time the poor relation in a field dominated by vision science, with the other senses lagging even further behind (Gallace and Spence, 2009; Gallace, 2013). Two things have happened to change this: firstly, from the 1980s, haptics has developed as a field in its own right; secondly, from the 1990s, there has been an accelerated interest in multisensory interactions. Here, we review the interactions and commonalities in visuo-haptic multisensory object processing, beginning with the capabilities and limits of haptic and visuo-haptic recognition. One way to facilitate recognition is to group like objects together: hence, we review recent work on the similarities between visual and haptic categorization and cross-modal transfer of category knowledge. Changes in orientation and size present a major challenge to within-modal object recognition. However, these obstacles seem to be absent in cross-modal recognition and we show that a shared representation underlies both cross-modal recognition and view-independence. We next compare visual and haptic representations from the point of view of individual differences in preferences for object or spatial imagery. A shared representation for vision and touch suggests shared neural processing and therefore we review a number of candidate brain regions, previously thought to be selective for visual aspects of object processing, which have subsequently been shown to be engaged by analogous haptic tasks. This reflects the growing consensus around the concept of a “metamodal” brain with a task-based organization and multisensory inputs, rather than organization around discrete unisensory inputs (Pascual-Leone and Hamilton, 2001; Lacey et al., 2009a; James et al., 2011). Finally, we draw these threads together and discuss the evidence

for a model of multisensory visuo-haptic object recognition in which representations are flexibly accessible by either top-down or bottom-up pathways depending on object familiarity and individual differences in imagery preference (Lacey et al., 2009a).

## HAPTIC AND VISUO-HAPTIC OBJECT RECOGNITION

The speed and accuracy of visual object recognition is well-established. Haptic recognition, albeit less well studied, is somewhat slower than visual recognition, but, at least for everyday objects, is still fairly fast and highly accurate with 96% correctly named: 68% in less than 3 s and 94% within 5 s (Klatzky et al., 1985); indeed, a “haptic glance” of less than 1 s suffices in some circumstances (Klatzky and Lederman, 1995). Longer response times in the study of Klatzky et al. (1985) likely reflect the time taken to explore some of the larger items such as a tennis racket or hairdryer. A remarkable fact about haptic processing is that it can be achieved with the feet as well as the hands, albeit more slowly and less accurately, with hand and foot performance being highly correlated across individuals (Lawson, 2014). Haptic identification proceeds, with increasing accuracy, from a “grasp and lift” stage that extracts basic low-level information about a variety of object properties to a series of hand movements that extract more precise information (Klatzky and Lederman, 1992). These hand movements, known as “exploratory procedures,” are property-specific, for example, lateral motion is used to assess texture and contour-following to precisely assess shape (Lederman and Klatzky, 1987). These properties differ in salience to haptic processing depending on the context: under neutral instructions, salience progressively decreases in this order: hardness > texture > shape; under instructions that emphasized haptic processing, the order changes to texture > shape > hardness

(Klatzky et al., 1987). Note that the saliency order under neutral instructions is reversed to shape > texture > hardness/size in simultaneous visual and haptic perception, and in haptic perception under instructions to use concurrent visual imagery (Klatzky et al., 1987).

Overall, cross-modal visuo-haptic object recognition, while fairly accurate, comes at a cost compared to within-modal recognition (e.g., Bushnell and Baxt, 1999; Casey and Newell, 2007; and see Lacey et al., 2007). Cross-modal performance is generally better when visual encoding is followed by haptic retrieval than the reverse (e.g., Jones, 1981; Streri and Molina, 1994; Lacey and Campbell, 2006). This asymmetry appears to be a consistent feature of visuo-haptic cross-modal memory but has generally received little attention (e.g., Easton et al., 1997a,b; Reales and Ballesteros, 1999; Nabeta and Kawahara, 2006). One explanation for cross-modal asymmetry might be that shape information is not encoded equally well by the visual and haptic systems, because of competition from other, more salient, modality-specific object properties. Thus, in the haptic-visual cross-modal condition it might be more difficult to encode shape because of the more salient hardness and texture information, as noted above. This effect might be suppressed by the use of concurrent visual imagery in which shape information, common to vision and touch, might be brought to the fore. We should note, however, that when vision and touch are employed simultaneously, properties that are differently weighted in these modalities may be optimally combined on the basis of maximum likelihood estimates (see Ernst and Banks, 2002; Helbig and Ernst, 2007; Helbig et al., 2012; Takahashi and Watt, 2014).

Another explanation for cross-modal asymmetry could be differences in visual and haptic memory capacity. Haptic working memory capacity appears to be limited and variable, and may therefore be more error-prone than visual working memory (Bliss and Hämaläinen, 2005). Alternatively, haptic representations may simply decay faster than visual representations. Rather than a progressive decline over time, the haptic decay function appears to occur entirely in a band of 15–30 s post-stimulus (Kiphart et al., 1992). Consistent with this, a more recent study showed no decline in performance at 15 s (Craddock and Lawson, 2010) although longer intervals were not tested. Haptic-visual performance might therefore be lower because by the time visual recognition is tested, haptically encoded representations have substantially decayed. However, other cross-modal memory studies show that delays up to 30 s (Garvill and Molander, 1973; Woods et al., 2004) or even a week (Pensky et al., 2008) did not affect haptic-visual recognition more than visual-haptic recognition. Thus, an explanation in terms of a simple function of haptic memory properties is likely insufficient.

Cross-modal asymmetry is observed even in very young infants where it is ascribed to constraints imposed by different stages of motor development (Streri and Molina, 1994). But this explanation is also unsatisfactory since the asymmetry persists into maturity (Easton et al., 1997a,b; Bushnell and Baxt, 1999; Lacey and Campbell, 2006). Interestingly, implicit memory does not appear to be affected: cross-modal priming is symmetric (Easton et al., 1997a,b; Reales and Ballesteros, 1999) although verbal encoding strategies may have played a

mitigating role in these studies. A recent study suggests that underlying neural activity is asymmetric between the two cross-modal conditions. Using a match-to-sample task, Kassuba et al. (2013) showed that bilateral lateral occipital complex (LOC), fusiform gyrus (FG), and anterior intraparietal sulcus (aIPS) selectively responded more strongly to crossmodal, compared to unimodal, object matching when haptic targets followed visual samples, and more strongly still when the haptic target and visual sample were congruent rather than incongruent; however, these regions showed no such increase for visual targets in either crossmodal or unimodal conditions. This asymmetric increase in activation in the visual-haptic condition may reflect multisensory binding of shape information and suggests that haptics – traditionally seen as the less reliable modality – has to integrate previously presented visual information more than vision has to integrate previous haptic information (Kassuba et al., 2013).

## OBJECT CATEGORIZATION

Categorization facilitates recognition and is critical for much of higher-order cognition (Graf, 2010); hitherto, the emphasis in terms of perceptual categorization has been almost exclusively on the visual, rather than the haptic, modality. More recently, however, a series of studies has systematically compared visual and haptic categorization. Using multi-dimensional scaling analysis, these studies showed that visual and haptic similarity ratings and categorization result in perceptual spaces [i.e., topological representations of the perceived (dis)similarity along a given dimension] that are highly congruent between modalities for novel 3-D objects (Cooke et al., 2007), more realistic 3-D shell-like objects (Gaißert et al., 2008, 2010, 2011) and for natural objects, i.e., actual seashells (Gaißert and Wallraven, 2012). This was so in both unisensory and bisensory conditions (Cooke et al., 2007) and whether 2-D visual objects were compared to haptic 3-D objects (Gaißert et al., 2008, 2010) or passive viewing of 2-D objects was compared to interactive viewing and active haptic exploration of 3-D objects, i.e., such that visual and haptic exploration were more similar (Gaißert et al., 2010). These highly similar visual and haptic perceptual spaces both showed high fidelity to the physical object space [i.e., a topological representation of the actual (dis)similarity along a given dimension; Gaißert et al., 2008, 2010], retaining the category structure (the ordinal adjacency relationships within the category, i.e., the actual progression in variation along a given dimension, for example from roughest to smoothest; Cooke et al., 2007). The isomorphism between perceptual (in either modality) and physical spaces was, furthermore, task-independent, whether simple similarity rating (Gaißert et al., 2008), unconstrained (free sorting), semi-constrained (making exactly three groups) or constrained (matching to a prototype object) categorization (Gaißert et al., 2011). As in vision, haptics also exhibits categorical perception, i.e., discriminability increases sharply when objects belong to different categories and decreases when they belong to the same category (Gaißert et al., 2012).

However, visual and haptic categorization are not entirely alike and, consistent with differential perceptual salience (Klatzky et al., 1987), object properties are differentially weighted depending on the modality, whether they are controlled parametrically (Cooke

et al., 2007) or vary naturally (Gaißert and Wallraven, 2012). Shape was more important than texture for visual categorization whereas in haptic and bisensory categorization, shape and texture were approximately equally weighted (Cooke et al., 2007), although in this study shape and texture varied in ways that were intuitive to vision and haptics (broadly, width for shape and smoothness for texture). Using specially manufactured shell-like objects, Gaißert et al. (2010) varied three complex shape parameters that were not intuitive to either modality. While visual and haptic perceptual spaces and the physical object space were all highly similar, the shape dimensions were weighted differently: symmetry was more important than convolutions for vision while the reverse was true for haptics; aperture-tip distance was the least important factor for both modalities (Gaißert et al., 2010). For natural objects – seashells – that varied naturally in a number of properties, similarity ratings and categorization were still driven by global and local shape parameters rather than size, texture, weight etc. (Gaißert and Wallraven, 2012).

These studies suggest a close connection between vision and haptics in terms of similarity mechanisms for categorization but do not necessarily imply a shared representation because of the differential weighting of object properties in each modality. Nonetheless, there is symmetric cross-modal transfer of category information following either visual or haptic category learning, even for complex novel 3-D objects, and furthermore this transfer generalizes to new objects from these categories (Yildirim and Jacobs, 2013). A recent study shows that not only does category membership transfer cross-modally, as shown by Yildirim and Jacobs (2013), but so does category structure (Wallraven et al., 2013), i.e., the ordinal relationships and category boundaries (see Cooke et al., 2007) transcend modality. Crossmodal transfer of category structure is interesting because the ordering of each item within the category is (at least in the studies reviewed here) perceptually driven; thus it may be that a shared multisensory representation underlies cross-modal categorization, as has been suggested for cross-modal recognition (Lacey et al., 2009a; Lacey and Sathian, 2011).

Of course, perceptual similarity is not the only basis for categorization (Smith et al., 1998) and neither vision nor haptics appear to naturally recover categories on alternative bases that are more abstract or semantic. For example, Haag (2011) used realistically textured models of familiar animals that retained real-life size relations, and required visual and haptic categorization on the basis of size (big/small in real life), domesticity (wild/domestic), and predation (carnivore/herbivore). Errors increased as the basis of categorization moved from concrete (size) to abstract (predation) and were consistently greater in haptics than vision (Haag, 2011). Similarly, neither vision nor haptics naturally recovered the taxonomic relationships between the natural seashells used by Gaißert and Wallraven (2012): participants distinguished between concrete categories such as whether the shells used were flat or convoluted, rather than between abstract categories such as gastropods (e.g., sea-snail) vs. bivalves (e.g., oyster). If biological relationships were recovered at all, this was mainly contingent on shape similarities, although vision was better than haptics in this respect (Gaißert and Wallraven, 2012) as it was for the abstract categories studied by Haag (2011).

## FACES: A SPECIAL CATEGORY

Faces are a special category of object that we encounter every day and at which we are especially expert, being able to differentiate large numbers of individuals (Maurer et al., 2002). We are also able to recognize faces under conditions that would impair recognition in other categories; for example, bad lighting or changes in viewpoint (Maurer et al., 2002) – though face recognition is impaired if the face is upside-down (Yin, 1969). An important distinction is made between configural and featural processing: the former refers to processing the spatial relationships between individual facial features as well as the shapes of the features themselves, while the latter refers to the piecemeal processing of individual face parts (Maurer et al., 2002; Dopjans et al., 2012). Although sighted humans obviously recognize faces almost exclusively through vision, live faces can also be identified haptically with high levels of accuracy (over 70%), whether they are learned through touch alone or using both vision and touch (Kilgour and Lederman, 2002). Interestingly, when participants had to haptically identify clay masks produced from live faces, accuracy was significantly lower than for live faces, suggesting that natural material cues and surface properties are important for haptic face recognition (Kilgour and Lederman, 2002). Visual experience may be necessary for haptic face recognition, since the congenitally blind were significantly less accurate than both the sighted and the late-blind (Wallraven and Dopjans, 2013). Nonetheless, haptic face recognition is not as good as visual recognition in the sighted either (Dopjans et al., 2012). This may be due to basic differences between visual and haptic processing. Haptic exploration of any object is almost exclusively sequential and serial (Lederman and Klatzky, 1987; Loomis et al., 1991) whilst visual processing is massively parallel (see Nassi and Callaway, 2009). In the context of face processing, therefore, haptics might be restricted to featural processing, in which individual features are processed independently and have to be assembled into a face context, which may account for lower haptic performance compared to visual configural encoding (Dopjans et al., 2012). When visual encoding was restricted, by using a participant-controlled moving window that only revealed a small portion of the face at a time, so that it was more like haptic sequential processing, visual and haptic performance were more equal (Dopjans et al., 2012), suggesting that any differences arise from different encoding strategies<sup>1</sup>.

Despite these various differences in performance, visual and haptic face processing do have common aspects. For example, consistent with the shared perceptual spaces discussed above (e.g., Gaißert et al., 2008, 2010, 2011; Gaißert and Wallraven, 2012), there is evidence for similar “face-spaces” for vision and touch in which, again, different properties carry different weights depending on the modality (Wallraven, 2014). The evidence for a face-inversion effect – better recognition when faces are upright than inverted, an effect not seen for non-face categories – is clear for vision but less so for haptics. Kilgour and Lederman (2006) showed a clear haptic inversion effect for faces compared to non-face stimuli, whereas Dopjans et al. (2012) found an inversion

<sup>1</sup>For discussions of configural versus featural visual face processing, see Peterson and Rhodes (2003).

effect for unrestricted visual, but not for haptic or restricted visual, face encoding. In “face adaptation,” a neutral face is perceived as having the opposite facial expression to a previously perceived face; for example, adaptation to a sad face leads to perception of a happy face upon subsequent presentation of a face with a neutral expression (e.g., Skinner and Benton, 2010). Such an effect is also seen in within-modal haptic adaptation to faces (Matsumiya, 2012) and transfers cross-modally both from vision to touch and *vice versa*, indicating that haptic face-related information and visual face processing share some common processing (Matsumiya, 2013).

Faces can also be recognized cross-modally between vision and touch (Kilgour and Lederman, 2002); this comes at a cost relative to within-modal recognition (Casey and Newell, 2007) although the cost decreases with familiarity (Casey and Newell, 2005). However, this disadvantage for cross-modal face recognition is unrelated to the encoding modality or to differences in encoding strategies, which suggests that, in contrast to object recognition (see below), vision and touch do not share a common face representation (Casey and Newell, 2007). On the other hand, visually presented faces disrupt identification of haptic faces when their facial expressions are incongruent and facilitate identification when they are congruent (Klatzky et al., 2011) which suggests a shared representation although response competition cannot be excluded as an explanation for these results. However, taken in conjunction with the finding that a visually prosopagnosic patient (i.e., a patient unable to recognize faces visually despite intact basic visual perception) was also unable to recognize faces haptically (Kilgour et al., 2004), a shared representation seems likely.

## OBSTACLES TO EFFICIENT RECOGNITION

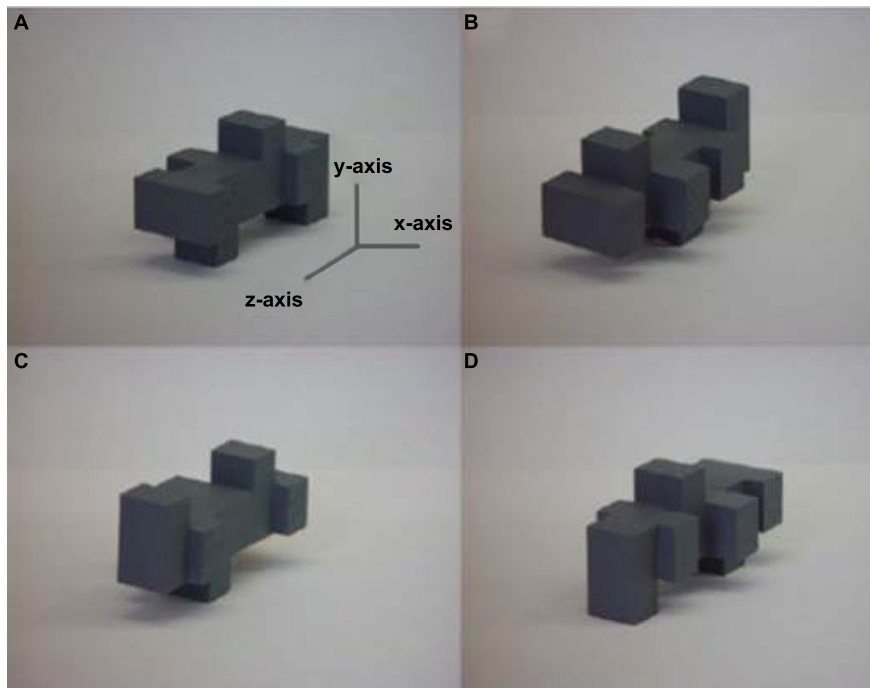
### VIEW-DEPENDENCE

A change in the orientation of an object changes the related sensory input, e.g., retinal pattern, such that recognition is potentially impaired; an important goal of sensory systems is therefore to achieve perceptual constancy so that objects can be recognized independently of such changes. Visual object recognition is considered view-dependent if rotating an object away from its original orientation impairs subsequent recognition and view-independent if not (reviewed in Peissig and Tarr, 2007). During haptic exploration, the hands can contact an object from different sides simultaneously: intuitively, therefore, one might expect information about several different “views” to be acquired at the same time and that haptic recognition would be view-independent. However, numerous studies have now shown that this intuition is not correct and that haptic object recognition is also view-dependent (Newell et al., 2001; Lacey et al., 2007, 2009b; Ueda and Saiki, 2007, 2012; Craddock and Lawson, 2008, 2010; Lawson, 2009, 2011). The factors underlying haptic view-dependence are not currently known: even unlimited exploration time and orientation cuing do not reduce view-dependence (Lawson, 2011). It is interesting to examine how vision and touch are affected by different types of rotation. Visual recognition is differentially impaired by changes in orientation depending on the axis around which an object is rotated (Gauthier et al., 2002; Lacey et al., 2007). Recognition is slower and less accurate when objects are rotated about the x- and y-axes, i.e., in depth (**Figure 1**), than

when rotated about the z-axis, i.e., in the picture plane, for both 2-D (Gauthier et al., 2002) and 3-D stimuli (Lacey et al., 2007). By contrast, haptic recognition is equally impaired by rotation about any axis (Lacey et al., 2007), suggesting that, although vision and haptics are both view-dependent, the basis for this is different in each modality. One possible explanation is that vision and haptics differ in whether or not a surface is occluded by rotation. In vision, a change in orientation can involve not only a transformation in perceptual shape but also occlusion of one or more surfaces – unless the observer physically changes position relative to the object (e.g., Pasqualotto et al., 2005; Pasqualotto and Newell, 2007). Compare, for example, **Figures 1A,C** – rotation about the x-axis means that the object is turned upside-down and that the former top surface becomes occluded. In haptic exploration, the hands are free to move over all surfaces of an object and to manipulate it into different orientations relative to the hand, thus in any given orientation, no surface is necessarily occluded, provided the object is small enough. If this is true, then no single axis of rotation should be more or less disruptive than another due to surface occlusion, so that haptic recognition only has to deal with a shape transformation. Further work is required to examine whether this explanation is, in fact, correct.

View-dependence mostly occurs when objects are unfamiliar. Increasing object familiarity reduces the disruptive effect of orientation changes and visual recognition tends to become view-independent (Tarr and Pinker, 1989; Bülthoff and Newell, 2006). An exception to this is when a familiar object is typically seen in one specific orientation known as a canonical view, for example the front view of a house (Palmer et al., 1981). View-independence may still occur for a limited range of orientations around the canonical view, but visual recognition is impaired for radically non-canonical views, for example, a teapot seen from directly above (Palmer et al., 1981; Tarr and Pinker, 1989; Bülthoff and Newell, 2006). Object familiarity also results in haptic view-independence and this remains so even where there is a change in the hand used to explore the object (Craddock and Lawson, 2009a). Haptic recognition also reverts to view-dependence for non-canonical orientations (Craddock and Lawson, 2008). However, vision and haptics differ in what constitutes a canonical view. The preferred view in vision is one in which the object is aligned at 45° to the observer (Palmer et al., 1981) while objects are generally aligned either parallel or orthogonal to the body midline in haptic canonical views (Woods et al., 2008). Canonical views may facilitate view-independent recognition either because they provide the most structural information about an object or because they most closely match a stored representation, but the end result is the same for both vision and haptics (Craddock and Lawson, 2008; Woods et al., 2008).

In contrast to within-modal recognition, visuo-haptic cross-modal recognition is view-independent even for unfamiliar objects that are highly similar (**Figure 1**), whether visual study is followed by haptic test or *vice versa* and whatever the axis of rotation (Lacey et al., 2007, 2010b; Ueda and Saiki, 2007, 2012). Haptic-visual, but not visual-haptic, cross-modal view-independence has been shown for familiar objects (Lawson, 2009). This asymmetry might



**FIGURE 1 | Example 3-D unfamiliar object shown (A) in the original orientation and rotated 180° about the (B) z-axis, (C) x-axis, and (D) y-axis: rotation about the x- and y-axes are rotations in depth, rotation about the z-axis is a rotation in the picture-plane.** Figure adapted from Lacey et al. (2007).

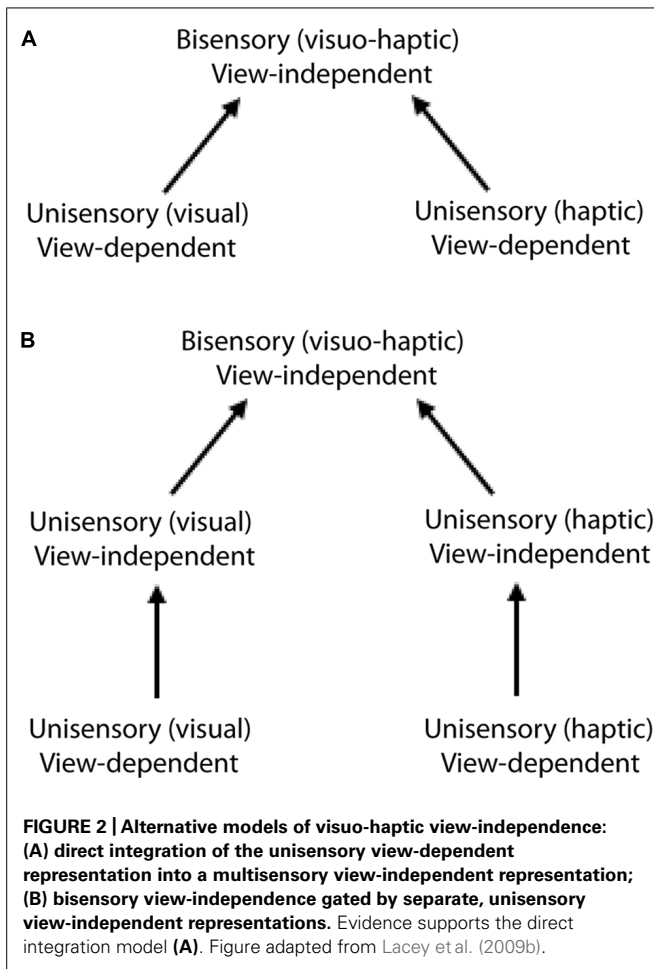
be due to the fact that the familiar objects used in this particular study were a mixture of scale models (e.g., bed, bath, and shark) and actual-size objects (e.g., jug, pencil); thus, some of these might have been more familiar visually than haptically, resulting in greater error when visually familiar objects had to be recognized by touch. Additional research on the potentially disruptive effects of differential familiarity is merited.

A strange finding is that knowledge of the test modality does not appear to help achieve view-independence. When participants knew the test modality, both visual and haptic within-modal recognition were view-dependent whereas cross-modal recognition was view-independent (Ueda and Saiki, 2007, 2012), but when the test modality was unknown both within- and cross-modal recognition were view-independent (Ueda and Saiki, 2007). At first glance this is puzzling: one would expect that knowledge of the test modality would confer an advantage. However, Ueda and Saiki (2012) showed that eye movements differed during encoding, with longer and more diffuse fixations when participants knew that they would be tested cross-modally (visual-haptic only) compared to within-modally. It is possible that, on the “principle of least commitment” (Marr, 1976), the same pattern of eye movements occurs when the test modality is not known (i.e., it is not possible to commit to an outcome), preserving as much information as possible and resulting in both within- and cross-modal view-independence. Further examination of eye movements during both cross-modal conditions would be valuable, as eye movements could serve as behavioral markers for the multisensory view-independent representation discussed next.

The simplest way in which cross-modal view-independence could arise is that the view-dependent visual and haptic unisensory representations are directly integrated into a view-independent multisensory representation (**Figure 2A**). An alternative explanation is that unisensory view-independence in vision and haptics is a precondition for cross-modal view-independence (**Figure 2B**). In a perceptual learning study, view-independence acquired by learning in one modality transferred completely and symmetrically to the other; thus, whether visual or haptic, within-modal view-independence relies on a single view-independent representation (Lacey et al., 2009b). Furthermore, both visual and haptic within-modal view-independence were acquired following cross-modal training (whether haptic-visual or visual-haptic); we therefore concluded that visuo-haptic view-independence is supported by a single multisensory representation that directly integrates the unisensory view-dependent representations (Lacey et al., 2009b; **Figure 2A**), similar to models that have been proposed for vision (Riesenhuber and Poggio, 1999). Thus, the same representation appears to support both cross-modal recognition and view-independence (whether within- or cross-modal).

#### SIZE-DEPENDENCE

In addition to achieving object constancy across orientation changes, the visual system also has to contend with variations in the size of the retinal image that arise from changes in object-observer distance: the same object can produce retinal images that vary in size depending on whether it is near to, or far from, the observer. Presumably, this is compensated by cues arising from



depth or motion perception, accounting for the fact that a change in size does not disrupt visual object identification (Biederman and Cooper, 1992; Uttl et al., 2007). However, size change does produce a cost in visual recognition for both unfamiliar (Jolicoeur, 1987) and familiar objects (Jolicoeur, 1987; Uttl et al., 2007). Interestingly, changes in retinal size due to movement of the observer result in better size-constancy than those due to movement of the object (Combe and Wexler, 2010).

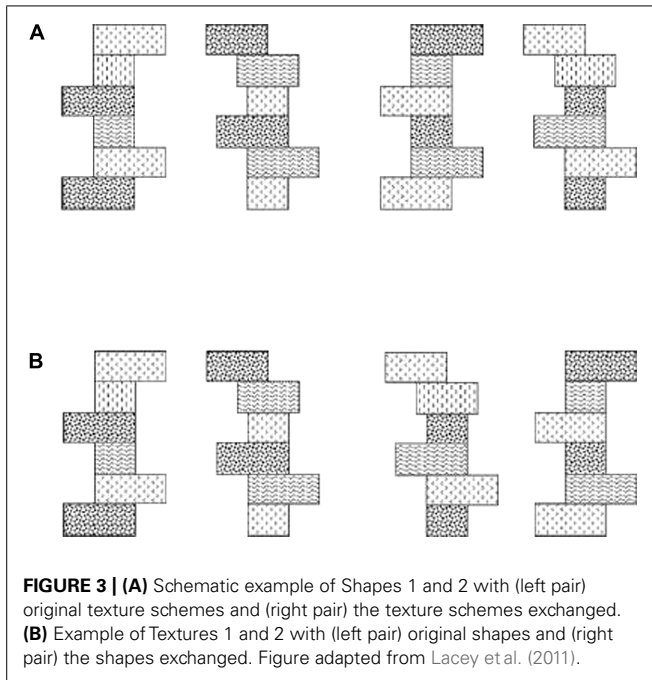
Haptic size perception requires integration of both cutaneous (contact area and force) and proprioceptive (finger spread and position) information at initial contact (Berryman et al., 2006). Neither gripping an object tighter, which increases contact area, nor enlarging the spread of the fingers leads us to perceive a change in size (Berryman et al., 2006). Thus, in contrast to vision where perceived size varies with distance, in touch, physical size is perceived directly, i.e., haptic size equals physical size. It is intriguing then, that haptic (Craddock and Lawson, 2009b,c) and cross-modal (Craddock and Lawson, 2009c) recognition are apparently size-dependent and this merits further investigation. Further research should address whether haptic representations store a canonical size for familiar objects (as has recently been proposed for visual representations, Konkle and Oliva, 2011), deviations from which could impair recognition, and whether

object constancy can be achieved across size changes in unfamiliar objects.

## REPRESENTATIONS AND INDIVIDUAL DIFFERENCES

A crucial question for object recognition is what information is contained in the mental representations that support it. Visual shape, color, and texture are processed in different cerebral cortical areas (Cant and Goodale, 2007; Cant et al., 2009) but these structural (shape) and surface (color, texture, etc.) properties are integrated in visual object representations (Nicholson and Humphrey, 2003). Changing the color of an object or its part-color combinations between study and test impaired shape recognition, while altering the background color against which objects were presented did not (Nicholson and Humphrey, 2003). This effect could therefore be isolated to the object representation, indicating that this contains both shape and color information (Nicholson and Humphrey, 2003). Visual and haptic within-modal object discrimination are similarly impaired by a change in surface texture (Lacey et al., 2010b), showing firstly that haptic representations also integrate structural and surface properties and secondly that information about surface properties in visual representations is not limited to modality-specific properties like color. In order to investigate whether surface properties are integrated into the multisensory representation underlying cross-modal object discrimination, we tested object discrimination across changes in orientation (thus requiring access to the view-independent multisensory representation discussed above), texture or both. In line with earlier findings (Lacey et al., 2007; Ueda and Saiki, 2007, 2012), cross-modal object discrimination was view-independent when texture did not change; but if texture did change, performance was reduced to chance levels, whether orientation also changed or not (Lacey et al., 2010b). However, some participants were more affected by the texture changes than others. We wondered whether this arose from individual differences in the nature of object representations, which can be conveniently indexed by preferences for different kinds of imagery.

Two kinds of visual imagery have been described: “object imagery” (involving pictorial images that are vivid and detailed, dealing with the literal appearance of objects in terms of shape, color, brightness, etc.) and “spatial imagery” (involving schematic images more concerned with the spatial relations of objects, their component parts, and spatial transformations; Kozhevnikov et al., 2002, 2005; Blajenkova et al., 2006). An experimentally important difference is that object imagery includes surface property information while spatial imagery does not. To establish whether object and spatial imagery differences occur in touch as well as vision, we required participants to discriminate shape across changes in texture, and texture across changes in shape (Figure 3), in both visual and haptic within-modal conditions. We found that spatial imagers could discriminate shape despite changes in texture but not *vice versa*, presumably because their images tend not to encode surface properties. By contrast, object imagers could discriminate texture despite changes in shape, but not the reverse (Lacey et al., 2011), indicating that texture, a surface property, is integrated into their shape representations. Importantly, visual and haptic performance was not significantly different on either



**FIGURE 3 | (A)** Schematic example of Shapes 1 and 2 with (left pair) original texture schemes and (right pair) the texture schemes exchanged. **(B)** Example of Textures 1 and 2 with (left pair) original shapes and (right pair) the shapes exchanged. Figure adapted from Lacey et al. (2011).

task and performance largely reflected both self-reports of imagery preference and scores on the Object and Spatial Imagery Questionnaire (OSIQ; Blajenkova et al., 2006). Thus, the object-spatial imagery continuum characterizes haptics as well as vision, and individual differences in imagery preference along this continuum affect the extent to which surface properties are integrated into object representations (Lacey et al., 2011). Further analysis of the texture-change condition in our earlier study (Lacey et al., 2010b) showed that performance was indeed related to imagery preference: both object and spatial imagers showed cross-modal view-independence but object imagers were impaired by texture changes whereas spatial imagers were not (Lacey et al., 2011). In addition, the extent of the impairment was correlated with OSIQ scores such that greater preference for object imagery was associated with greater impairment by texture changes; surface properties are therefore likely only integrated into the multisensory representation by object imagers (Lacey et al., 2011). Moreover, spatial imagery preference correlated with the accuracy of cross-modal object recognition (Lacey et al., 2007). It appears, then, that the multisensory representation has some features that are stable across individuals, like view-independence, and some that vary across individuals, such as integration of surface property information and individual differences in imagery preference.

### THE NEURAL BASIS OF VISUO-HAPTIC OBJECT PROCESSING SEGREGATED VENTRAL “WHAT” AND DORSAL “WHERE/HOW” PATHWAYS

At the macro-level, visual object processing divides along a ventral pathway concerned with object identity and perception for recognition, and a dorsal pathway dealing with object location and perception for action, e.g., reaching and grasping, (Ungerleider and Mishkin, 1982; Goodale and Milner, 1992). Similar ventral

and dorsal pathways have been proposed for the auditory (e.g., De Santis et al., 2007a) and somatosensory domains (Dijkerman and de Haan, 2007), with divergence of the “what” and “where/how” pathways in a similar timeframe (~200 ms after stimulus onset) (De Santis et al., 2007a,b), and thus are probably common aspects of functional architecture across modalities.

In the case of touch, an early functional magnetic resonance (fMRI) study found that haptic object recognition activated frontal cortical areas as well as inferior parietal cortex, while a haptic object location task activated superior parietal regions (Reed et al., 2005). A later study from our laboratory (Sathian et al., 2011) compared perception of haptic texture and location, reasoning that texture would be a better marker of haptic object identity, given the salience of texture to touch (Klatzky et al., 1987). This study found that, while both visual and haptic location judgments involved a similar dorsal pathway comprising large sectors of the IPS and frontal eye fields (FEFs) bilaterally, haptic texture perception engaged extensive areas of the parietal operculum (OP), which contains higher-order (i.e., non-primary), ventral regions of somatosensory cortex. In addition, shared cortical processing of texture across vision and touch was found in parts of extrastriate (i.e., non-primary) visual cortex and ventral premotor cortex (Sathian et al., 2011). For both texture and location, several of these bisensory areas showed correlations of activation magnitude between the visual and haptic tasks, indicating some commonality of cortical processing across modalities (Sathian et al., 2011). Another group extended these findings by showing that early visual cortex showed activation magnitudes that not only scaled with the interdot spacing of dot-patterns, but were also modulated by the presence of matching haptic input (Eck et al., 2013).

### MULTISENSORY PROCESSING OF OBJECT SHAPE

Cortical areas in both the ventral and dorsal pathways previously identified as specialized for various aspects of visual processing are also functionally involved during the corresponding haptic tasks (for reviews see Amedi et al., 2005; Sathian and Lacey, 2007; Lacey and Sathian, 2011). In the human visual pathway even early visual areas (which project to both dorsal and ventral streams) have been found to respond to changes in haptic shape, suggesting that haptic shape perception might involve the entire ventral stream (Snow et al., 2014). If true, this might reflect cortical pathways between primary somatosensory and visual cortices previously demonstrated in the macaque (Négyessy et al., 2006); however, as with other studies (see below), it is not possible to exclude visual imagery as an explanation for the findings of Snow et al. (2014). The majority of research on visuo-haptic processing of object shape has concentrated on higher-level visual areas, in particular the LOC, an object-selective region in the ventral visual pathway (Malach et al., 1995), a sub-region of which also responds selectively to objects in both vision and touch (Amedi et al., 2001, 2002; Stilla and Sathian, 2008). The LOC responds to both haptic 3-D (Amedi et al., 2001; Zhang et al., 2004; Stilla and Sathian, 2008) and tactile 2-D stimuli (Stoesz et al., 2003; Prather et al., 2004) but does not respond during auditory object recognition cued by object-specific sounds (Amedi et al., 2002). However, when participants listened to the impact sounds made by rods and balls made of either metal or wood and categorized these

sounds by the shape of the object that made them, the material of the object, or by using all the acoustic information available, the LOC was more activated when these sounds were categorized by shape than by material (James et al., 2011). Here again though, participants could have solved this matching task using visual imagery: we return to the potential role of visual imagery in a later section.

The LOC does, however, respond to auditory shape information created by a visual-auditory sensory substitution device (Amedi et al., 2007) using a specific algorithm to convert visual information into an auditory stream or “soundscape” in which the visual horizontal axis is represented by auditory duration and stereo panning, the visual vertical axis by variations in tone frequency, and pixel brightness by variations in tone loudness. Although it requires extensive training, both sighted and blind humans can learn to recognize objects by extracting shape information from such soundscapes (Amedi et al., 2007). However, the LOC only responds to soundscapes created according to the algorithm – and which therefore represent shape in a principled way – and not when participants learn soundscapes that are merely arbitrarily associated with particular objects (Amedi et al., 2007). Thus, the LOC can be regarded as processing geometric shape information independently of the sensory modality used to acquire it.

Apart from the LOC, multisensory (visuo-haptic) responses have also been observed in several parietal regions: in particular, the aIPS is involved in perception of both the shape and location of objects, with co-activation of the LOC for shape and the FEF for location (Stilla and Sathian, 2008; Sathian et al., 2011; see also Saito et al., 2003). The postcentral sulcus (PCS; Stilla and Sathian, 2008), corresponding to Brodmann’s area 2 of primary somatosensory cortex (S1; Grefkes et al., 2001), also shows visuo-haptic shape-selectivity. This area is normally considered exclusively somatosensory but the bisensory responses observed by Stilla and Sathian (2008) are consistent with earlier neurophysiological studies that suggested visual responsiveness in parts of S1 (Iriki et al., 1996; Zhou and Fuster, 1997).

Multisensory responses in the LOC and elsewhere might reflect visuo-haptic integration in neurons that process both visual and haptic input; alternatively, they might arise from separate inputs to discrete but interdigitated unisensory neuronal populations. Tal and Amedi (2009) sought to distinguish between these using fMRI adaptation (fMR-A). This technique utilizes the repetition suppression effect, i.e., when the same stimulus is repeated, the blood-oxygen level dependent (BOLD) signal is attenuated. Since repetition suppression can be observed in single neurons, fMR-A can reveal neuronal selectivity profiles (see Grill-Spector et al., 2006; Kregelberg et al., 2006 for reviews). When stimuli that had been presented visually were presented again haptically, there was a robust cross-modal adaptation effect not only in the LOC and the aIPS, but also in bilateral precentral sulcus (preCS) corresponding to ventral premotor cortex, and the right anterior insula, suggesting that these areas were integrating multisensory inputs at the neuronal level. However, a separate preCS site and posterior parts of the IPS did not show cross-modal adaptation, suggesting that their multisensory responses arise from separate unisensory populations. Because fMR-A effects may not necessarily reflect neuronal

selectivity (Mur et al., 2010), it will be necessary to confirm the findings of Tal and Amedi (2009) with converging evidence using other methods.

It is critical to determine whether haptic or tactile involvement in supposedly visual cortical areas is functionally relevant, i.e., whether it is actually necessary for task performance. Although research along these lines is still relatively sparse, two lines of evidence indicate that this is indeed the case. Firstly, case studies indicate that the LOC is necessary for both haptic and visual shape perception. A lesion to the left occipito-temporal cortex, which likely included the LOC, resulted in both tactile and visual agnosia even though somatosensory cortex and basic somatosensory function were intact (Feinberg et al., 1986). Another patient with bilateral LOC lesions was unable to learn new objects either visually or haptically (James et al., 2006b). These case studies are consistent with the existence of a shared multisensory representation in the LOC.

Transcranial magnetic stimulation (TMS) is a technique used to temporarily deactivate specific, functionally defined, cortical areas, i.e., to create “virtual lesions” (Sack, 2006). TMS over a parieto-occipital region previously shown to be active during tactile grating orientation discrimination (Sathian et al., 1997) interfered with performance of this task (Zangaladze et al., 1999) indicating that it was functionally, rather than epiphenomenally, involved. This area is the probable human homolog of macaque area V6 (Pitzalis et al., 2006). Repetitive TMS (rTMS) over the left LOC impaired visual object, but not scene, categorization (Mullin and Steeves, 2011), similarly suggesting that this area is necessary for object processing. rTMS over the left aIPS impaired visual-haptic, but not haptic-visual, shape matching using the right hand (Buelte et al., 2008), but shape matching with the left hand during rTMS over the right aIPS was unaffected in either cross-modal condition. The reason for this discrepancy is unclear, and emphasizes that the precise roles of the IPS and LOC in multisensory shape processing have yet to be fully worked out.

### CATEGORY-SPECIFIC REPRESENTATIONS

There has been rather limited neural study of cross-modal category-selective representations. Using multivoxel pattern analysis of fMRI data, Pietrini et al. (2004) demonstrated that selectivity for particular categories of man-made objects was correlated across vision and touch in a region of inferotemporal cortex. In the case of face perception, fMRI studies, in contrast to the behavioral studies reviewed above, tend to favor separate, rather than shared representations. For example, visual and haptic face-selectivity in ventral and inferior temporal cortex are in largely separate voxel populations (Pietrini et al., 2004). Haptic face recognition activates the left FG, whereas visual face recognition activates the right FG (Kilgour et al., 2005); furthermore, activity in the left FG increases during haptic processing of familiar, compared to unfamiliar, faces while the right FG remains relatively inactive (James et al., 2006a). A further difference in FG face responses is that imagery of visually presented faces activates the left FG more than the right FG (Ishai et al., 2002)<sup>2</sup>; this raises the possibility that haptic face perception

<sup>2</sup>Note that, although these studies mainly refer to the fusiform gyrus, this is not the only cortical region involved in face processing, nor are faces necessarily the only



involves visual imagery mechanisms. Although one study found that haptic face recognition ability and imagery vividness ratings were uncorrelated (Kilgour and Lederman, 2002), the implication of visual imagery in haptic face perception is very consonant with our findings in haptic shape perception discussed below (Deshpande et al., 2010; Lacey et al., 2010a) especially as vividness ratings do not particularly index imagery ability (reviewed in Lacey and Lawson, 2013). Further studies are needed to resolve the neural basis of multisensory face perception, and its differences from multisensory object perception.

### VIEW- AND SIZE-INDEPENDENCE

The cortical locus of the multisensory view-independent representation is currently not known. Evidence for visual view-independence in the LOC is mixed: as might be expected, unfamiliar objects produce view-dependent LOC responses (Gauthier et al., 2002) and familiar objects produce view-independent responses (Valyear et al., 2006; Eger et al., 2008a; Pourtois et al., 2009). By contrast, one study found view-dependence in the LOC even for familiar objects, although in this study there was position-independence (Grill-Spector et al., 1999), whereas another found view-independence for both familiar and unfamiliar objects (James et al., 2002a). A recent TMS study of 2-D shape suggests that the LOC is functionally involved in view-independent recognition (Silvanto et al., 2010) but only two rotations, 20 and 70°, were tested and TMS effects were only seen for the 20° rotation; further work is required to substantiate this finding. Responses in the FG are also variable with the left FG less sensitive to orientation changes than the right FG (Andresen et al., 2009; Harvey and Burgund, 2012). A study of face viewpoint-selectivity showed a gradient of decreasing orientation sensitivity, from view-dependence in early visual cortex to partial view-independence in later areas including LOC (Axelrod and Yovel, 2012); this sensitivity gradient may also apply to non-face objects.

Various parietal regions show visual view-dependent responses, e.g., the IPS (James et al., 2002a) and a parieto-occipital area (Valyear et al., 2006). Superior parietal cortex is view-dependent during mental rotation but not visual object recognition (Gauthier et al., 2002; Wilson and Farah, 2006). As these regions are in the dorsal pathway, concerned with object location and perception for action, view-dependent responses in these regions are not surprising (Ungerleider and Mishkin, 1982; Goodale and Milner, 1992). Actions such as reaching and grasping adapt to changes in object orientation and consistent with this, lateral parieto-occipital cortex shows view-dependent responses for graspable, but not for non-graspable objects (Rice et al., 2007).

To date, we are not aware of neuroimaging studies of haptic or cross-modal processing of stimuli across changes in orientation. James et al. (2002b) varied object orientation, but this study concentrated on haptic-to-visual priming rather than the cross-modal response to same vs. different orientations *per se*. Additionally, there is much work to be done on the effect of orientation changes when shape information is derived from the auditory soundscapes

produced by sensory substitution devices (SSDs) and also when the options for haptically interacting with an object are altered by a change in orientation. Similarly, there is no neuroimaging work on haptic and multisensory processing of stimuli across changes in size. However, visual size-independence has been consistently observed in the LOC (Grill-Spector et al., 1999; Ewbank et al., 2005; Eger et al., 2008a,b), with anterior regions showing more size-independence than posterior regions (Sawamura et al., 2005; Eger et al., 2008b).

### A MODEL OF VISUO-HAPTIC MULTISENSORY OBJECT REPRESENTATION

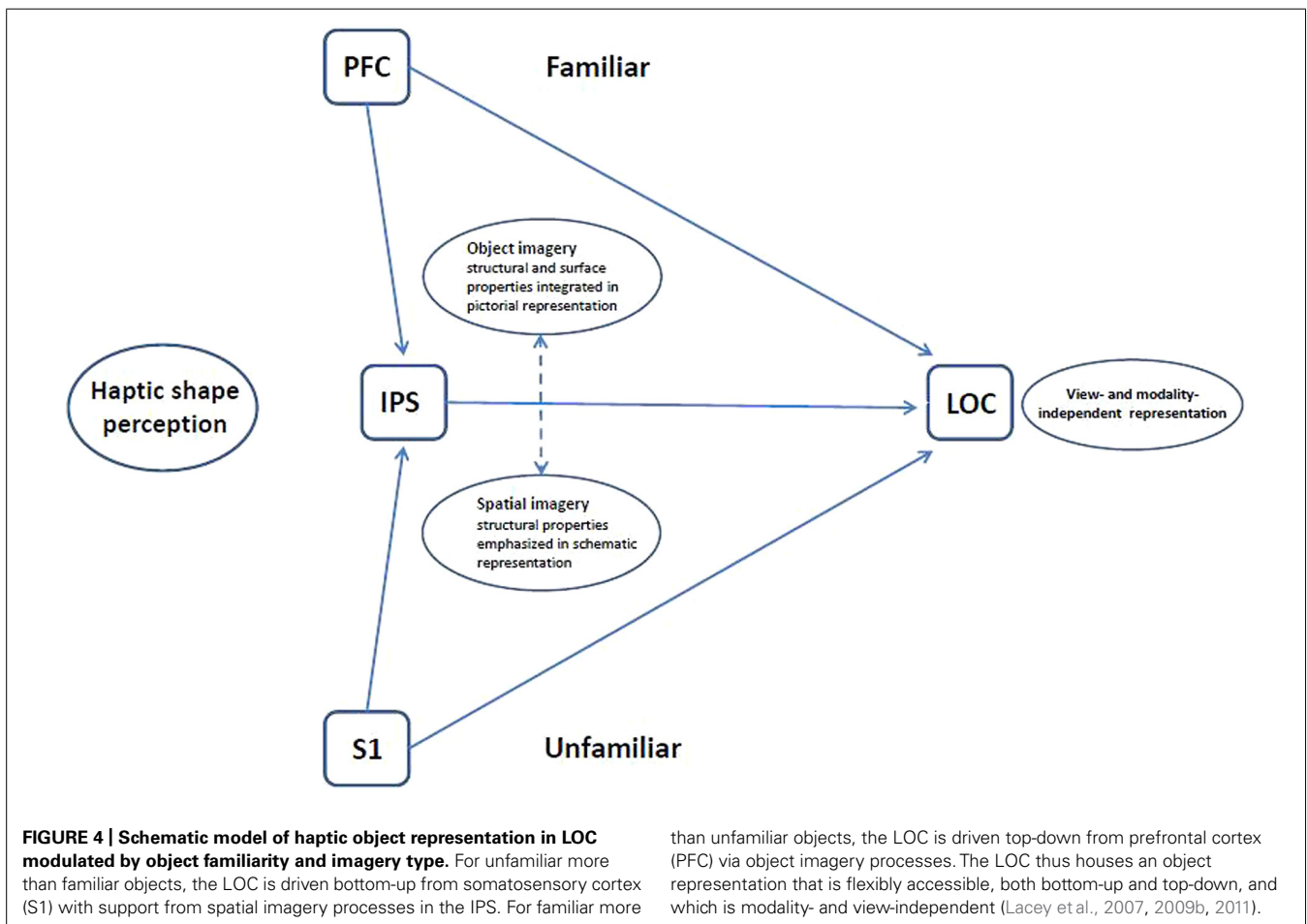
Haptic activation of the LOC might arise from direct somatosensory input. Activity in somatosensory cortex propagates to the LOC as early as 150 ms after stimulus onset during tactile discrimination of simple shapes, a timeframe consistent with “bottom-up” projections to LOC (Lucan et al., 2010; Adhikari et al., 2014). Similarly, in a tactile microspatial discrimination task, LOC activity was consistent with feedforward propagation in a beta-band oscillatory network (Adhikari et al., 2014). In addition, a patient with bilateral ventral occipito-temporal lesions, but with sparing of the dorsal part of the LOC that likely included the multisensory sub-region, showed visual agnosia but intact haptic object recognition (Allen and Humphreys, 2009). Haptic object recognition was associated with activation of the intact dorsal part of the LOC, suggesting that somatosensory input could directly activate this region (Allen and Humphreys, 2009).

Alternatively, haptic perception might evoke visual imagery of the felt object resulting in “top-down” activation of the LOC (Sathian et al., 1997) and consistent with this hypothesis, many studies show LOC activity during visual imagery. During auditorily cued mental imagery of familiar object shape, both blind and sighted participants show left LOC activation, where shape information would arise mainly from haptic experience for the blind and mainly from visual experience for the sighted (De Volder et al., 2001). The left LOC is also active when geometric and material object properties are retrieved from memory (Newman et al., 2005) and haptic shape-selective activation magnitudes in the right LOC were highly correlated with ratings of visual imagery vividness (Zhang et al., 2004). A counter-argument is that imagery plays a relatively minor role because LOC activity was substantially lower during visual imagery compared to haptic shape perception (Amedi et al., 2001). However, this study could not verify that participants engaged in imagery throughout the imaging session, so that lower imagery-related activity might have resulted from non-compliance (or irregular compliance) with the task. It has also been argued that visual imagery cannot explain haptically evoked LOC activity because early- as well as late-blind individuals show shape-related LOC activation via both touch (reviewed in Pascual-Leone et al., 2005; Sathian, 2005; Sathian and Lacey, 2007) and hearing using SSDs (Arno et al., 2001; Renier et al., 2004, 2005; Amedi et al., 2007). But this argument, while true for the early blind, does not rule out a visual imagery explanation in the sighted, given the extensive evidence for cross-modal plasticity following visual deprivation (reviewed in Pascual-Leone et al., 2005; Sathian, 2005; Sathian and Lacey, 2007).

category processed in that, or other, regions; this issue remains controversial (see Harel et al., 2013, for a review).

In this section we describe a model of visuo-haptic multisensory object representation (Lacey et al., 2009a) and review the evidence for this model from studies designed to explicitly test the visual imagery hypothesis discussed above (Deshpande et al., 2010; Lacey et al., 2010a, 2014). In this model, object representations in the LOC can be flexibly accessed either bottom-up or top-down, depending on object familiarity, and independently of the input modality. There is no stored representation for unfamiliar objects so that during haptic recognition, an unfamiliar object has to be explored in its entirety in order to compute global shape and to relate component parts to one another. This, we propose, occurs in a bottom-up pathway from somatosensory cortex to the LOC, with involvement of the IPS in computing part relationships and thence global shape, facilitated by spatial imagery processes. For familiar objects, global shape can be inferred more easily, perhaps from distinctive features or one diagnostic part, and we suggest that haptic exploration rapidly acquires enough information to trigger a stored visual image and generate a hypothesis about its identity, as has been proposed for vision (e.g., Bar, 2007). This occurs in a top-down pathway from prefrontal cortex to LOC, involving primarily object imagery processes (though spatial imagery may still have a role in processing familiar objects, for example, in view-independent recognition).

We tested this model using analyses of inter-task correlations of activation magnitude between visual object imagery and haptic shape perception (Lacey et al., 2010a) and analyses of effective connectivity (Deshpande et al., 2010), reasoning that reliance on similar processes across tasks would lead to correlations of activation magnitude across participants, as well as similar patterns of effective connectivity across tasks. In contrast to previous studies, we ensured that participants engaged in visual imagery throughout each scan by using an object imagery task and recording responses. Participants also performed a haptic shape discrimination task using either familiar or unfamiliar objects. We found that object familiarity modulated inter-task correlations as predicted by our model. There were eleven regions common to visual object imagery and haptic perception of familiar shape, six of which (including bilateral LOC) showed inter-task correlations of activation magnitude. By contrast, object imagery and haptic perception of unfamiliar shape shared only four regions, only one of which (an IPS region) showed an inter-task correlation (Lacey et al., 2010a). More recently, we examined the relation between haptic shape perception and spatial imagery, using a spatial imagery task in which participants memorized a  $4 \times 4$  lettered grid and, in response to auditory letter strings, constructed novel shapes within the imagined grid from component parts (Lacey et al., 2014); the haptic shape tasks were the same



as in Lacey et al. (2010a). Contrary to the model, relatively few regions showed inter-task correlations between spatial imagery and haptic perception of either familiar or unfamiliar shape, with parietal foci featuring in both sets of correlations. This suggests that spatial imagery is relevant to haptic shape perception regardless of object familiarity, whereas our earlier finding suggested that object imagery is more strongly associated with haptic perception of familiar, than unfamiliar, shape (Lacey et al., 2010a). However, it is also possible that the parietal foci showing inter-task correlations between spatial imagery and haptic shape perception reflect spatial processing more generally, rather than spatial imagery *per se* (Lacey et al., 2014; and see Jäncke et al., 2001), or generic imagery processes, e.g., image generation, common to both object and spatial imagery (Lacey et al., 2014; and see Mechelli et al., 2004).

In our study of spatial imagery (Lacey et al., 2014), we also conducted effective connectivity analyses, based on the inferred neuronal activity derived from deconvolving the hemodynamic response out of the observed BOLD signals (Sathian et al., 2013). In order to make direct comparisons between the neural networks underlying object and spatial imagery in haptic shape perception, we re-analyzed our earlier data (Deshpande et al., 2010) using the newer effective connectivity methods. These analyses supported the broad architecture of the model, showing that the spatial imagery network shared much more commonality with the network associated with unfamiliar, compared to familiar, shape perception, while the object imagery network shared much more commonality with familiar, than unfamiliar, shape perception (Lacey et al., 2014). More specifically, the model proposes that the component parts of an unfamiliar object are explored in their entirety and assembled into a representation of global shape via spatial imagery processes (Lacey et al., 2009a). Consistent with this, in the parts of the network that were common to spatial imagery and unfamiliar haptic shape perception, the LOC is driven by parietal foci, with complex cross-talk between posterior parietal and somatosensory foci. These findings fit with the notion of bottom-up pathways from somatosensory cortex and a role for cortex in and around the IPS in spatial imagery (Lacey et al., 2014). The IPS and somatosensory interactions were absent from the sparse network that was shared by spatial imagery and haptic perception of familiar shape. By contrast, the relationship between object imagery and familiar shape perception is characterized by top-down pathways from prefrontal areas reflecting the involvement of object imagery, according to our model (Lacey et al., 2009a). The re-analyzed data supported this, showing the LOC driven bilaterally by the left inferior frontal gyrus in the network shared by object imagery and haptic perception of familiar shape, while these pathways were absent from the extremely sparse network common to object imagery and unfamiliar haptic shape perception (Lacey et al., 2014).

**Figure 4** shows the current version of our model for haptic shape perception in which the LOC is driven bottom-up from primary somatosensory cortex as well as top-down via object imagery processes from prefrontal cortex, with additional input from the IPS involving spatial imagery processes. We propose that the bottom-up route is more important for haptic perception of

unfamiliar than familiar objects, whereas the converse is true of the top-down route – more important for haptic perception of familiar than unfamiliar objects. It will be interesting to explore the impact of individual preferences for object vs. spatial imagery on these processes and paths.

## SUMMARY

The research reviewed here illustrates how deeply interconnected the visual and haptic modalities are in object processing, from highly similar and transferable perceptual spaces underlying categorization, through shared representations in cross-modal and view-independent recognition and commonalities in imagery preferences, to multisensory neural substrates and complex interactions between bottom-up and top-down processes as well as between object and spatial imagery. Much, however, remains to be done in order to provide a detailed account of visuo-haptic multisensory behavior and its underlying mechanisms and how this understanding can be put to use, for example in the service of neurorehabilitation, particularly for those with sensory deprivation of various sorts.

## ACKNOWLEDGMENTS

Support to K. Sathian from the National Eye Institute at the NIH, the National Science Foundation, and the Veterans Administration is gratefully acknowledged.

## REFERENCES

- Adhikari, B. M., Sathian, K., Epstein, C. M., Lamichhane, B., and Dhamala, M. (2014). Oscillatory activity in neocortical networks during tactile discrimination near the limit of spatial acuity. *Neuroimage* 91, 300–310. doi: 10.1016/j.neuroimage.2014.01.007
- Allen, H. A., and Humphreys, G. W. (2009). Direct tactile stimulation of dorsal occipito-temporal cortex in a visual agnostic. *Curr. Biol.* 19, 1044–1049. doi: 10.1016/j.cub.2009.04.057
- Amedi, A., Jacobson, G., Hendler, T., Malach, R., and Zohary, E. (2002). Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cereb. Cortex* 12, 1202–1212. doi: 10.1093/cercor/12.11.1202
- Amedi, A., Malach, R., Hendler, T., Peled, S., and Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nat. Neurosci.* 4, 324–330. doi: 10.1038/85201
- Amedi, A., Stern, W. M., Camprodon, J. A., Bempohl, F., Merabet, L., Rotman, S., et al. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nat. Neurosci.* 10, 687–689. doi: 10.1038/nm1912
- Amedi, A., von Kriegstein, K., van Atteveldt, N. M., Beauchamp, M. S., and Naumer, M. J. (2005). Functional imaging of human crossmodal identification and object recognition. *Exp. Brain Res.* 166, 559–571. doi: 10.1007/s00221-005-2396-5
- Andresen, D. R., Vinberg, J., and Grill-Spector, K. (2009). The representation of object viewpoint in human visual cortex. *Neuroimage* 45, 522–536. doi: 10.1016/j.neuroimage.2008.11.009
- Arno, P., De Volder, A. G., Vanlierde, A., Wanet-Defalque, M.-C., Streel, E., Robert, A., et al. (2001). Occipital activation by pattern recognition in the early blind using auditory substitution for vision. *Neuroimage* 13, 632–645. doi: 10.1006/nimg.2000.0731
- Axelrod, V., and Yovel, G. (2012). Hierarchical processing of face viewpoint in human visual cortex. *J. Neurosci.* 32, 2442–2452. doi: 10.1523/JNEUROSCI.4770-11.2012
- Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends Cogn. Sci.* 11, 280–289. doi: 10.1016/j.tics.2007.05.005
- Berryman, L. J., Yau, J. M., and Hsiao, S. S. (2006). Representation of object size in the somatosensory system. *J. Neurophysiol.* 96, 27–39. doi: 10.1152/jn.01190.2005
- Biederman, I., and Cooper, E. E. (1992). Size invariance in visual object priming. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 121–133. doi: 10.1037/0096-1523.18.1.121

- Blajenkova, O., Kozhevnikov, M., and Motes, M. A. (2006). Object-spatial imagery: a new self-report imagery questionnaire. *Appl. Cogn. Psychol.* 20, 239–263. doi: 10.1002/acp.1182
- Bliss, I., and Hämäläinen, H. (2005). Different working memory capacity in normal young adults for visual and tactile letter recognition task. *Scand. J. Psychol.* 46, 247–251. doi: 10.1111/j.1467-9450.2005.00454.x
- Buelte, D., Meister, I. G., Staedtgen, M., Dambeck, N., Sparing, R., Grefkes, C., et al. (2008). The role of the anterior intraparietal sulcus in crossmodal processing of object features in humans: an rTMS study. *Brain Res.* 1217, 110–118. doi: 10.1016/j.brainres.2008.03.075
- Bülthoff, I., and Newell, F. N. (2006). The role of familiarity in the recognition of static and dynamic objects. *Prog. Brain Res.* 154, 315–325. doi: 10.1016/S0079-6123(06)54017-8
- Bushnell, E. W., and Baxt, C. (1999). Children's haptic and cross-modal recognition with familiar and unfamiliar objects. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 1867–1881. doi: 10.1037/0096-1523.25.6.1867
- Cant, J. S., Arnott, S. R., and Goodale, M. A. (2009). fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. *Exp. Brain Res.* 192, 391–405. doi: 10.1007/s00221-008-1573-8
- Cant, J. S., and Goodale, M. A. (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb. Cortex* 17, 713–731. doi: 10.1093/cercor/bhk022
- Casey, S. J., and Newell, F. N. (2005). The role of long-term and short-term familiarity in visual and haptic face recognition. *Exp. Brain Res.* 166, 583–591. doi: 10.1007/s00221-005-2398-3
- Casey, S. J., and Newell, F. N. (2007). Are representations of unfamiliar faces independent of encoding modality? *Neuropsychologia* 45, 506–513. doi: 10.1016/j.neuropsychologia.2006.02.011
- Combe, E., and Wexler, M. (2010). Observer movement and size constancy. *Psychol. Sci.* 21, 667–675. doi: 10.1177/0956797610367753
- Cooke, T., Jäkel, F., Wallraven, C., and Bülthoff, H. H. (2007). Multimodal similarity and categorization of novel, three-dimensional objects. *Neuropsychologia* 45, 484–495. doi: 10.1016/j.neuropsychologia.2006.02.009
- Craddock, M., and Lawson, R. (2008). Repetition priming and the haptic recognition of familiar and unfamiliar objects. *Percept. Psychophys.* 70, 1350–1365. doi: 10.3758/PP.70.7.1350
- Craddock, M., and Lawson, R. (2009a). Do left and right matter for haptic recognition of familiar objects? *Perception* 38, 1355–1376. doi: 10.1068/p6312
- Craddock, M., and Lawson, R. (2009b). The effect of size changes on haptic object recognition. *Atten. Percept. Psychophys.* 71, 910–923. doi: 10.3758/APP.71.4.910
- Craddock, M., and Lawson, R. (2009c). Size-sensitive perceptual representations underlie visual and haptic object recognition. *PLoS ONE* 4:e8009. doi: 10.1371/journal.pone.0008009
- Craddock, M., and Lawson, R. (2010). The effects of temporal delay and orientation on haptic object recognition. *Atten. Percept. Psychophys.* 72, 1975–1980. doi: 10.3758/APP.72.7.1975
- De Santis, L., Clarke, S., and Murray, M. M. (2007a). Automatic and intrinsic auditory 'what' and 'where' processing in humans revealed by electrical neuroimaging. *Cereb. Cortex* 17, 9–17. doi: 10.1093/cercor/bhj119
- De Santis, L., Spierer, L., Clarke, S., and Murray, M. M. (2007b). Getting in touch: segregated somatosensory what and where pathways in humans revealed by electrical neuroimaging. *Neuroimage* 37, 890–903. doi: 10.1016/j.neuroimage.2007.05.052
- Deshpande, G., Hu, X., Lacey, S., Stilla, R., and Sathian, K. (2010). Object familiarity modulates effective connectivity during haptic shape perception. *Neuroimage* 49, 1991–2000. doi: 10.1016/j.neuroimage.2009.08.052
- De Volder, A. G., Toyama, H., Kimura, Y., Kiyosawa, M., Nakano, H., Vanlierde, A., et al. (2001). Auditory triggered mental imagery of shape involves visual association areas in early blind humans. *Neuroimage* 14, 129–139. doi: 10.1006/nimg.2001.0782
- Dijkerman, H. C., and de Haan, E. H. F. (2007). Somatosensory processes subserving perception and action. *Behav. Brain Sci.* 30, 189–239. doi: 10.1017/S0140525X07001392
- Dopjans, L., Bülthoff, H. H., and Wallraven, C. (2012). Serial exploration of faces: comparing vision and touch. *J. Vis.* 12, 6. doi: 10.1167/12.1.6
- Easton, R. D., Greene, A. J., and Srinivas, K. (1997a). Transfer between vision and haptics: memory for 2-D patterns and 3-D objects. *Psychon. Bull. Rev.* 4, 403–410. doi: 10.3758/BF03210801
- Easton, R. D., Srinivas, K., and Greene, A. J. (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. doi: 10.1037/0278-7393.23.1.153
- Eck, J., Kaas, A. L., and Goebel, R. (2013). Crossmodal interactions of haptic and visual texture information in early sensory cortex. *Neuroimage* 75, 123–135. doi: 10.1016/j.neuroimage.2013.02.075
- Eger, E., Ashburner, J., Haynes, J.-D., Dolan, R. J., and Rees, G. (2008a). fMRI activity patterns in human LOC carry information about object exemplars within category. *J. Cogn. Neurosci.* 20, 356–370. doi: 10.1162/jocn.2008.20019
- Eger, E., Kell, C. A., and Kleinschmidt, A. (2008b). Graded size-sensitivity of object-exemplar-evoked activity patterns within human LOC regions. *J. Neurophysiol.* 100, 2038–2047. doi: 10.1152/jn.90305.2008
- Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433. doi: 10.1038/415429a
- Ewbank, M. P., Schluppeck, D., and Andrews, T. J. (2005). fMR-adaptation reveals a distributed representation of inanimate objects and places in human visual cortex. *Neuroimage* 28, 268–279. doi: 10.1016/j.neuroimage.2005.06.036
- Feinberg, T. E., Rothi, L. J., and Heilman, K. M. (1986). Multimodal agnosia after unilateral left hemisphere lesion. *Neurology* 36, 864–867. doi: 10.1212/WNL.36.6.864
- Gaißert, N., Bülthoff, H. H., and Wallraven, C. (2011). Similarity and categorization: from vision to touch. *Acta Psychol.* 138, 219–230. doi: 10.1016/j.actpsy.2011.06.007
- Gaißert, N., and Wallraven, C. (2012). Categorizing natural objects: a comparison of the visual and haptic modalities. *Exp. Brain Res.* 216, 123–134. doi: 10.1007/s00221-011-2916-4
- Gaißert, N., Wallraven, C., and Bülthoff, H. H. (2008). "Analyzing perceptual representations of complex, parametrically-defined shapes using MDS," in *Proceedings of the Sixth International Conference EuroHaptics 2008, Lecture Notes in Computer Science – Haptics: Perception, Devices and Scenarios, Vol. 5024* (Heidelberg: Springer Berlin Heidelberg), 265–274.
- Gaißert, N., Wallraven, C., and Bülthoff, H. H. (2010). Visual and haptic perceptual spaces show high similarity in humans. *J. Vis.* 10, 2. doi: 10.1167/10.11.2
- Gaißert, N., Waterkamp, S., Fleming, R. W., and Bülthoff, I. (2012). Haptic categorical perception of shape. *PLoS ONE* 7:e43062. doi: 10.1371/journal.pone.0043062
- Gallace, A. (2013). "Somesthetic mental imagery," in *Multisensory Imagery*, eds S. Lacey and R. Lawson (New York: Springer), 29–50. doi: 10.1007/978-1-4614-5879-1\_3
- Gallace, A., and Spence, C. (2009). The cognitive and neural correlates of tactile memory. *Psychol. Bull.* 135, 380–406. doi: 10.1037/a0015325
- Garvill, J., and Molander, B. (1973). Effects of standard modality, comparison modality and retention interval on matching of form. *Scand. J. Psychol.* 14, 203–206. doi: 10.1111/j.1467-9450.1973.tb00111.x
- Gauthier, I., Hayward, W. G., Tarr, M. J., Anderson, A. W., Skudlarski, P., and Gore, J. C. (2002). BOLD activity during mental rotation and view-dependent object recognition. *Neuron* 34, 161–171. doi: 10.1016/S0896-6273(02)00622-0
- Goodale, M. A., and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25. doi: 10.1016/0166-2236(92)90344-8
- Graf, M. (2010). "Categorization and object shape," in *Towards a Theory of Thinking: Building Blocks for a Conceptual Framework*, eds B. M. Glatzeder, V. Goel, and A. von Müller (Berlin: Springer-Verlag), 73–101. doi: 10.1007/978-3-642-03129-8\_6
- Grefkes, C., Geyer, S., Schormann, T., Roland, P., and Zilles, K. (2001). Human somatosensory area 2: observer-independent cytoarchitectonic mapping, interindividual variability, and population map. *Neuroimage* 14, 617–631. doi: 10.1006/nimg.2001.0858
- Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23. doi: 10.1016/j.tics.2005.11.006
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., and Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24, 187–203. doi: 10.1016/S0896-6273(00)80832-6
- Haag, S. (2011). Effects of vision and haptics on categorizing common objects. *Cogn. Process.* 12, 33–39. doi: 10.1007/s10339-010-0369-5

- Harel, A., Kravitz, D., and Baker, C. I. (2013). Beyond perceptual expertise: revisiting the neural substrates of expert object recognition. *Front. Hum. Neurosci.* 7:885. doi: 10.3389/fnhum.2013.00885
- Harvey, D. Y., and Burgund, E. D. (2012). Neural adaptation across viewpoint and exemplar in fusiform cortex. *Brain Cogn.* 80, 33–44. doi: 10.1016/j.bandc.2012.04.009
- Helbig, H. B., and Ernst, M. O. (2007). Optimal integration of shape information from vision and touch. *Exp. Brain Res.* 179, 595–606. doi: 10.1007/s00221-006-0814-y
- Helbig, H. B., Ernst, M. O., Ricciardi, E., Pietrini, P., Thielscher, A., Mayer, K. M., et al. (2012). The neural mechanisms of reliability weighted integration of shape information from vision and touch. *Neuroimage* 60, 1063–1072. doi: 10.1016/j.neuroimage.2011.09.072
- Iriki, A., Tanaka, M., and Iwamura, Y. (1996). Attention-induced neuronal activity in the monkey somatosensory cortex revealed by pupillometrics. *Neurosci. Res.* 25, 173–181. doi: 10.1016/0168-0102(96)01043-7
- Ishai, A., Haxby, J. V., and Ungerleider, L. G. (2002). Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *Neuroimage* 17, 1729–1741. doi: 10.1006/nimg.2002.1330
- James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., and Goodale, M. A. (2002a). Differential effects of view on object-driven activation in dorsal and ventral streams. *Neuron* 35, 793–801. doi: 10.1016/S0896-6273(02)00803-6
- James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S., and Goodale, M. A. (2002b). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia* 40, 1706–1714. doi: 10.1016/S0028-3932(02)00017-9
- James, T. W., Servos, P., Kilgour, A. R., Huh, E., and Lederman, S. (2006a). The influence of familiarity on brain activation during haptic exploration of 3-D facemasks. *Neurosci. Lett.* 397, 269–273. doi: 10.1016/j.neulet.2005.12.052
- James, T. W., James, K. H., Humphrey, G. K., and Goodale, M. A. (2006b). “Do visual and tactile object representations share the same neural substrate?” in *Touch and Blindness: Psychology and Neuroscience*, eds M. A. Heller and S. Ballesteros (Mahwah, NJ: Lawrence Erlbaum Associates), 139–155.
- James, T. W., Stevenson, R. W., Kim, S., VanDerKlok, R. M., and James, K. H. (2011). Shape from sound: evidence for a shape operator in the lateral occipital cortex. *Neuropsychologia* 49, 1807–1815. doi: 10.1016/j.neuropsychologia.2011.03.004
- Jäncke, L., Kleinschmidt, A., Mirzazade, S., Shah, N. J., and Freund, H.-J. (2001). The role of the inferior parietal cortex in linking the tactile perception and manual construction of object shapes. *Cereb. Cortex* 11, 114–121. doi: 10.1093/cercor/11.2.114
- Jolicoeur, P. (1987). A size-congruency effect in memory for visual shape. *Mem. Cogn.* 15, 531–543. doi: 10.3758/BF03198388
- Jones, B. (1981). “The developmental significance of cross-modal matching,” in *Intersensory Perception and Sensory Integration*, eds R. D. Walk and H. L. Jr. Pick, (New York: Plenum Press), 108–136.
- Kassuba, T., Klinge, C., Höllig, C., Röder, B., and Siebner, H. R. (2013). Vision holds a greater share in visuo-haptic object recognition than touch. *Neuroimage* 65, 59–68. doi: 10.1016/j.neuroimage.2012.09.054
- Kilgour, A. R., de Gelder, B., and Lederman, S. (2004). Haptic face recognition and prosopagnosia. *Neuropsychologia* 42, 707–712. doi: 10.1016/j.neuropsychologia.2003.11.021
- Kilgour, A. R., Kitada, R., Servos, P., James, T. W., and Lederman, S. J. (2005). Haptic face identification activates ventral occipital and temporal areas: an fMRI study. *Brain Cogn.* 59, 246–257. doi: 10.1016/j.bandc.2005.07.004
- Kilgour, A. R., and Lederman, S. (2002). Face recognition by hand. *Percept. Psychophys.* 64, 339–352. doi: 10.3758/BF03194708
- Kilgour, A. R., and Lederman, S. (2006). A haptic face-inversion effect. *Perception* 35, 921–931. doi: 10.1068/p5341
- Kiphart, M. J., Hughes, J. L., Simmons, J. P., and Cross, H. A. (1992). Short-term haptic memory for complex objects. *Bull. Psychon. Soc.* 30, 212–214. doi: 10.3758/BF03330444
- Klatzky, R. L., Abramowicz, A., Hamilton, C., and Lederman, S. J. (2011). Irrelevant visual faces influence haptic identification of facial expressions of emotion. *Atten. Percept. Psychophys.* 73, 521–530. doi: 10.3758/s13414-010-0038-x
- Klatzky, R. J., and Lederman, S. J. (1992). Stages of manual exploration in haptic object identification. *Percept. Psychophys.* 52, 661–670. doi: 10.3758/BF03211702
- Klatzky, R. L., and Lederman, S. J. (1995). Identifying objects from a haptic glance. *Percept. Psychophys.* 57, 1111–1123. doi: 10.3758/BF03208368
- Klatzky, R. L., Lederman, S. J., and Metzger, V. A. (1985). Identifying objects by touch: an ‘expert system’. *Percept. Psychophys.* 37, 299–302. doi: 10.3758/BF03211351
- Klatzky, R. L., Lederman, S. J., and 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. doi: 10.1037/0096-3445.116.4.356
- Konkle, T., and Oliva, A. (2011). Canonical visual size for real-world objects. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 23–37. doi: 10.1037/a0020413
- Kozhevnikov, M., Hegarty, M., and Mayer, R. E. (2002). Revising the visualiser-verbaliser dimension: evidence for two types of visualisers. *Cogn. Instr.* 20, 47–77. doi: 10.1207/S1532690XCI2001\_3
- Kozhevnikov, M., Kosslyn, S. M., and Shephard, J. (2005). Spatial versus object visualisers: a new characterisation of cognitive style. *Mem. Cogn.* 33, 710–726. doi: 10.3758/BF03195337
- Krekelberg, B., Boynton, G. M., and van Wezel, R. J. A. (2006). Adaptation: from single cells to BOLD signals. *Trends Neurosci.* 29, 250–256. doi: 10.1016/j.tins.2006.02.008
- Lacey, S., and Campbell, C. (2006). Mental representation in visual/haptic cross-modal memory: evidence from interference effects. *Q. J. Exp. Psychol.* 59, 361–376. doi: 10.1080/17470210500173232
- Lacey, S., Flueckiger, P., Stilla, R., Lava, M., and Sathian, K. (2010a). Object familiarity modulates the relationship between visual object imagery and haptic shape perception. *Neuroimage* 49, 1977–1990. doi: 10.1016/j.neuroimage.2009.10.081
- Lacey, S., Hall, J., and Sathian, K. (2010b). Are surface properties integrated into visuo-haptic object representations? *Eur. J. Neurosci.* 31, 1882–1888. doi: 10.1111/j.1460-9568.2010.07204.x
- Lacey, S., and Lawson, R. (2013). “Imagery questionnaires: vividness and beyond,” in *Multisensory Imagery*, eds S. Lacey and R. Lawson (New York: Springer), 271–282.
- Lacey, S., Lin, J. B., and Sathian, K. (2011). Object and spatial imagery dimensions in visuo-haptic representations. *Exp. Brain Res.* 213, 267–273. doi: 10.1007/s00221-011-2623-1
- Lacey, S., Peters, A., and Sathian, K. (2007). Cross-modal object representation is viewpoint-independent. *PLoS ONE* 2:e890. doi: 10.1371/journal.pone0000890
- Lacey, S., and Sathian, K. (2011). Multisensory object representation: insights from studies of vision and touch. *Prog. Brain Res.* 191, 165–176. doi: 10.1016/B978-0-444-53752-2.00006-0
- Lacey, S., Stilla, R., Sreenivasan, K., Deshpande, G., and Sathian, K. (2014). Spatial imagery in haptic shape perception. *Neuropsychologia* doi: 10.1016/j.neuropsychologia.2014.05.008 [Epub ahead of print].
- Lacey, S., Tal, N., Amedi, A., and Sathian, K. (2009a). A putative model of multisensory object representation. *Brain Topogr.* 21, 269–274. doi: 10.1007/s10548-009-0087-4
- Lacey, S., Pappas, M., Kreps, A., Lee, K., and Sathian, K. (2009b). Perceptual learning of view-independence in visuo-haptic object representations. *Exp. Brain Res.* 198, 329–337. doi: 10.1007/s00221-009-1856-8
- Lawson, R. (2009). A comparison of the effects of depth rotation on visual and haptic three-dimensional object recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 911–930. doi: 10.1037/a0015025
- Lawson, R. (2011). An investigation into the cause of orientation-sensitivity in haptic object recognition. *Seeing Perceiving* 24, 293–314. doi: 10.1163/187847511X579052
- Lawson, R. (2014). Recognizing familiar objects by hand and foot: haptic shape perception generalizes to inputs from unusual locations and untrained body parts. *Atten. Percept. Psychophys.* 76, 541–558. doi: 10.3758/s13414-013-0559-1
- Lederman, S. J., and Klatzky, R. L. (1987). Hand movements: a window into haptic object recognition. *Cogn. Psychol.* 19, 342–368. doi: 10.1016/0010-0285(87)90008-9
- Loomis, J., Klatzky, R. L., and Lederman, S. J. (1991). Similarity of tactual and visual picture recognition with limited field of view. *Perception* 20, 167–177. doi: 10.1068/p200167
- Lucan, J. N., Foxe, J. J., Gomez-Ramirez, M., Sathian, K., and Molholm, S. (2010). Tactile shape discrimination recruits human lateral occipital complex during early perceptual processing. *Hum. Brain Mapp.* 31, 1813–1821. doi: 10.1002/hbm.20983
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance

- imaging in human occipital cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 8135–8139. doi: 10.1073/pnas.92.18.8135
- Marr, D. (1976). Early processing of visual information. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 275, 483–524. doi: 10.1098/rstb.1976.0090
- Matsumiya, K. (2012). Haptic face aftereffect. *Iperception* 3, 97–100.
- Matsumiya, K. (2013). Seeing a haptically explored face: visual facial-expression aftereffect from haptic adaptation to a face. *Psychol. Sci.* 24, 2088–2098. doi: 10.1177/0956797613486981
- Maurer, D., Le Grand, R., and Mondloch, C. J. (2002). The many faces of configural processing. *Trends Cogn. Sci.* 6, 255–260. doi: 10.1016/S1364-6613(02)01903-4
- Mechelli, A., Price, C. J., Friston, K. J., and Ishai, A. (2004). Where bottom-up meets top-down: neuronal interactions during perception and imagery. *Cereb. Cortex* 14, 1256–1265. doi: 10.1093/cercor/bhh087
- Mullin, C. R., and Steeves, J. K. E. (2011). TMS to the lateral occipital cortex disrupts object processing but facilitates scene processing. *J. Cogn. Neurosci.* 23, 4174–4184. doi: 10.1162/jocn\_a\_00095
- Mur, M., Ruff, D. A., Bodurka, J., Bandettini, P. A., and Kriegeskorte, N. (2010). Face-identity change activation outside the face system: “release from adaptation” may not always indicate neuronal selectivity. *Cereb. Cortex* 20, 2027–2042. doi: 10.1093/cercor/bhp272
- Nabeta, T., and Kawahara, J. (2006). Congruency effect of presentation modality on false recognition of haptic and visual objects. *Memory* 14, 307–315. doi: 10.1080/09658210500277398
- Nassi, J. J., and Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nat. Rev. Neurosci.* 10, 360–372. doi: 10.1038/nrn2619
- Négyessy, L., Nepusz, T., Kocsis, L., and Bazsó, F. (2006). Prediction of the main cortical areas and connections involved in the tactile function of the visual cortex by network analysis. *Eur. J. Neurosci.* 23, 1919–1930. doi: 10.1111/j.1460-9568.2006.04678.x
- Newell, F. N., Ernst, M. O., Tjan, B. S., and Bühlhoff, H. H. (2001). View dependence in visual and haptic object recognition. *Psychol. Sci.* 12, 37–42. doi: 10.1111/1467-9280.00307
- Newman, S. D., Klatzky, R. L., Lederman, S. J., and Just, M. A. (2005). Imagining material versus geometric properties of objects: an fMRI study. *Cogn. Brain Res.* 23, 235–246. doi: 10.1016/j.cogbrainres.2004.10.020
- Nicholson, K. G., and Humphrey, G. K. (2003). The effect of colour congruency on shape discriminations of novel objects. *Perception* 32, 339–353. doi: 10.1002/hbm.20983
- Palmer, S. E., Rosch, E., and Chase, P. (1981). “Canonical perspective and the perception of objects,” in *Attention and Performance IX*, eds J. Long and A. Baddeley (Hillsdale, NJ: Lawrence Erlbaum Associates), 135–151.
- Pascual-Leone, A., Amedi, A., Fregni, F., and Merabet, L. B. (2005). The plastic human brain. *Annu. Rev. Neurosci.* 28, 377–401. doi: 10.1146/annurev.neuro.27.070203.144216
- Pascual-Leone, A., and Hamilton, R. H. (2001). The metamodal organization of the brain. *Prog. Brain Res.* 134, 427–445. doi: 10.1016/S0079-6123(01)34028-1
- Pasqualotto, A., Finucane, C., and Newell, F. N. (2005). Visual and haptic representations of scenes are updated with observer movement. *Exp. Brain Res.* 166, 481–488. doi: 10.1007/s00221-005-2388-5
- Pasqualotto, A., and Newell, F. N. (2007). The role of visual experience on the representation and updating of novel haptic scenes. *Brain Cogn.* 65, 184–194. doi: 10.1016/j.bandc.2007.07.009
- Peissig, J. J., and Tarr, M. J. (2007). Visual object recognition: do we know more now than we did 20 years ago? *Annu. Rev. Psychol.* 58, 75–96. doi: 10.1146/annurev.psych.58.102904.190114
- Pensky, A. E. C., Johnson, K. A., Haag, S., and Homa, D. (2008). Delayed memory for visual-haptic exploration of objects. *Psychon. Bull. Rev.* 15, 574–580. doi: 10.3758/PBR.15.3.574
- Peterson, M. A., and Rhodes, G. (2003). *Perception of Faces, Objects, and Scenes: Analytic and Holistic Processes*. Oxford: Oxford University Press.
- Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W.-H. C., Cohen, L., et al. (2004). Beyond sensory images: object-based representation in the human ventral pathway. *Proc. Natl. Acad. Sci. U.S.A.* 101, 5658–5663. doi: 10.1073/pnas.0400707101
- Pitzalis, S., Galletti, C., Huang, R. S., Patria, F., Committeri, G., Galati, G., et al. (2006). Wide-field retinotopy defines human cortical visual area V6. *J. Neurosci.* 26, 7962–7973. doi: 10.1523/JNEUROSCI.0178-06.2006
- Pourtois, G., Schwarz, S., Spiridon, M., Martuzzi, R., and Vuilleumier, P. (2009). Object representations for multiple visual categories overlap in lateral occipital and medial fusiform cortex. *Cereb. Cortex* 19, 1806–1819. doi: 10.1093/cercor/bhn210
- Prather, S. C., Votaw, J. R., and Sathian, K. (2004). Task-specific recruitment of dorsal and ventral visual areas during tactile perception. *Neuropsychologia* 42, 1079–1087. doi: 10.1016/j.neuropsychologia.2003.12.013
- Reales, J. M., and Ballesteros, S. (1999). Implicit and explicit memory for visual and haptic objects: cross-modal priming depends on structural descriptions. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 644–663. doi: 10.1037/0278-7393.25.3.644
- Reed, C. L., Klatzky, R. L., and Halgren, E. (2005). What vs. where in touch: an fMRI study. *Neuroimage* 25, 718–726. doi: 10.1016/j.neuroimage.2004.11.044
- Renier, L., Collignon, O., Tranduy, D., Poirier, C., Vanlierde, A., Veraart, C., et al. (2004). Visual cortex activation in early blind and sighted subjects using an auditory visual substitution device to perceive depth. *Neuroimage* 22, S1.
- Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., et al. (2005). Cross modal activation of visual cortex during depth perception using auditory substitution of vision. *Neuroimage* 26, 573–580. doi: 10.1016/j.neuroimage.2005.01.047
- Rice, N. J., Valyear, K. F., Goodale, M. A., Milner, A. D., and Culham, J. C. (2007). Orientation sensitivity to graspable objects: an fMRI adaptation study. *Neuroimage* 36, T87–T93. doi: 10.1016/j.neuroimage.2007.03.032
- Riesenhuber, M., and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2, 1019–1025. doi: 10.1038/14819
- Sack, A. T. (2006). Transcranial magnetic stimulation, causal structure-function mapping and networks of functional relevance. *Curr. Opin. Neurobiol.* 16, 593–599. doi: 10.1016/j.conb.2006.06.016
- Saito, D. N., Okada, T., Morita, Y., Yonekura, Y., and Sadato, N. (2003). Tactile-visual cross-modal shape matching: a functional MRI study. *Cogn. Brain Res.* 17, 14–25. doi: 10.1016/S0926-6410(03)00076-4
- Sathian, K. (2005). Visual cortical activity during tactile perception in the sighted and the visually deprived. *Dev. Psychobiol.* 46, 279–286. doi: 10.1002/dev.20056
- Sathian, K., Deshpande, G., and Stilla, R. (2013). Neural changes with tactile learning reflect decision-level reweighting of perceptual readout. *J. Neurosci.* 33, 5387–5398. doi: 10.1523/JNEUROSCI.3482-12.2013
- Sathian, K., and Lacey, S. (2007). Journeying beyond classical somatosensory cortex. *Can. J. Exp. Psychol.* 61, 254–264. doi: 10.1037/cjep2007026
- Sathian, K., Lacey, S., Stilla, R., Gibson, G., Deshpande, G., Hu, X., et al. (2011). Dual pathways for somatosensory information processing. *Neuroimage* 57, 462–475. doi: 10.1016/j.neuroimage.2011.05.001
- Sathian, K., Zangaladze, A., Hoffman, J. M., and Grafton, S. T. (1997). Feeling with the mind’s eye. *Neuroreport* 8, 3877–3881. doi: 10.1097/00001756-199712220-00008
- Sawamura, H., Georgieva, S., Vogels, R., Vanduffel, W., and Orban, G. A. (2005). Using functional magnetic resonance imaging to assess adaptation and size invariance of shape processing by humans and monkeys. *J. Neurosci.* 25, 4294–4306. doi: 10.1523/JNEUROSCI.0377-05.2005
- Silvanto, J., Schwarzkopf, D. S., Gilaie-Dotan, S., and Rees, G. (2010). Differing causal roles for lateral occipital complex and occipital face area in invariant shape recognition. *Eur. J. Neurosci.* 32, 165–171. doi: 10.1111/j.1460-9568.2010.07278.x
- Skinner, A. L., and Benton, C. P. (2010). Anti-expression aftereffects reveal prototype-referenced coding of facial expressions. *Psychol. Sci.* 21, 1248–1253. doi: 10.1177/0956797610380702
- Smith, E. E., Patalano, A. L., and Jonides, J. (1998). Alternative strategies of categorization. *Cognition* 65, 167–196. doi: 10.1016/S0010-0277(97)00043-7
- Snow, J. C., Strother, L., and Humphreys, G. W. (2014). Haptic shape processing in visual cortex. *J. Cogn. Neurosci.* 26, 1154–1167. doi: 10.1162/jocn\_a\_00548
- Stilla, R., and Sathian, K. (2008). Selective visuo-haptic processing of shape and texture. *Hum. Brain Mapp.* 29, 1123–1138. doi: 10.1002/hbm.20456
- Stoesz, N., Zhang, M., Weisser, V. D., Prather, S. C., Mao, H., and 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. doi: 10.1016/S0167-8760(03)00123-5
- Streri, A., and Molina, M. (1994). “Constraints on intermodal transfer between touch and vision in infancy,” in *The Development of Intersensory Perception: Comparative Perspectives*, eds D. J. Lewkowicz and R. Lickliter (Hove: Lawrence Erlbaum Associates), 285–307.

- Takahashi, C., and Watt, S. J. (2014). Visual-haptic integration with pliers and tongs: signal 'weights' take account of changes in haptic sensitivity caused by different tools. *Front. Psychol.* 5:109. doi: 10.3389/fpsyg.2014.00109
- Tal, N., and 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. doi: 10.1007/s00221-009-1949-4
- Tarr, M. J., and Pinker, S. (1989). Mental rotation and orientation dependence in shape recognition. *Cogn. Psychol.* 21, 233–282. doi: 10.1016/0010-0285(89)90009-1
- Ueda, Y., and Saiki, J. (2007). Viewpoint independence in visual and haptic object recognition. *Jpn. J. Psychon. Sci.* 26, 11–19.
- Ueda, Y., and Saiki, J. (2012). Characteristics of eye movements in 3-D object learning: comparison between within-modal and cross-modal object recognition. *Perception* 41, 1289–1298. doi: 10.1068/p7257
- Ungerleider, L. G., and Mishkin, M. (1982). "Two cortical visual systems," in *Analysis of Visual Behavior*, eds D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield (Cambridge, MA: MIT Press), 549–586.
- Uttl, B., Graf, P., and Siegenthaler, A. L. (2007). Influence of object size on baseline identification, priming, and explicit memory. *Scand. J. Psychol.* 48, 281–288. doi: 10.1111/j.1467-9450.2007.00571.x
- Valyear, K. F., Culham, J. C., Sharif, N., Westwood, D., and Goodale, M. A. (2006). A double dissociation between sensitivity to changes in object identity and object orientation in the ventral and dorsal streams: a human fMRI study. *Neuropsychologia* 44, 218–228. doi: 10.1016/j.neuropsychologia.2005.05.004
- Wallraven, C. (2014). Touching on face space: comparing visual, and haptic processing of face shapes. *Psychon. Bull. Rev.* doi: 10.3758/s13423-013-0577-y [Epub ahead of print].
- Wallraven, C., Bühlhoff, H. H., Waterkamp, S., van Dam, L., and Gaißert, N. (2013). The eyes grasp, the hands see: metric category knowledge transfers between vision and touch. *Psychon. Bull. Rev.* doi: 10.3758/s13423-013-0563-4 [Epub ahead of print].
- Wallraven, C., and Dopjans, L. (2013). Visual experience is necessary for efficient haptic face recognition. *Neuroreport* 24, 254–258. doi: 10.1097/WNR.0b013e32835f00c0
- Wilson, K., and Farah, M. J. (2006). Distinct patterns of viewpoint-dependent BOLD activity during common-object recognition and mental rotation. *Perception* 35, 1351–1366. doi: 10.1068/p5571
- Woods, A. T., Moore, A., and Newell, F. N. (2008). Canonical views in haptic object perception. *Perception* 37, 1867–1878. doi: 10.1068/p6038
- Woods, A. T., O'Modhrain, S., and Newell, F. N. (2004). The effect of temporal delay and spatial differences on crossmodal object recognition. *Cogn. Affect. Behav. Neurosci.* 4, 260–269. doi: 10.3758/CABN.4.2.260
- Yildirim, I., and Jacobs, R. A. (2013). Transfer of object category knowledge across visual and haptic modalities: experimental and computational studies. *Cognition* 126, 135–148. doi: 10.1016/j.cognition.2012.08.005
- Yin, R. K. (1969). Looking at upside-down faces. *J. Exp. Psychol.* 81, 141–145. doi: 10.1037/h0027474
- Zangaladze, A., Epstein, C. M., Grafton, S. T., and Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature* 401, 587–590. doi: 10.1038/44139
- Zhang, M., Weisser, V. D., Stilla, R., Prather, S. C., and Sathian, K. (2004). Multisensory cortical processing of object shape and its relation to mental imagery. *Cogn. Affect. Behav. Neurosci.* 4, 251–259. doi: 10.3758/CABN.4.2.251
- Zhou, Y.-D., and Fuster, J. M. (1997). Neuronal activity of somatosensory cortex in a cross-modal (visuo-haptic) memory task. *Exp. Brain Res.* 116, 551–555. doi: 10.1007/PL00005783

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 February 2014; accepted: 23 June 2014; published online: 17 July 2014.

Citation: Lacey S and Sathian K (2014) Visuo-haptic multisensory object recognition, categorization, and representation. *Front. Psychol.* 5:730. doi: 10.3389/fpsyg.2014.00730

This article was submitted to Perception Science, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Lacey and Sathian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.