"Sapienza" University of Rome

Department of Psychology

Ph.D. course in "Psychology and Cognitive Science"

XXIX cycle

The Role of Visual Processing in Haptic Representation

Recognition tasks with novel 3D objects

Tutor: Prof. Marta Olivetti Belardinelli

Co-Tutors: Prof. Franco Delogu

Prof. Daniele Nardi

Ph.D. Student

Mariangela Gravina

Index

1. Introduction

By now, all of us know that in perceiving and recognizing everyday objects we use different senses combined together (multisensory process). However, vision has always received almost the whole attention in research fields leaving the other sensory modalities scarcely explored (Gallace and Spence, 2009). This probably because, as it is easy to imagine, we mainly explore and recognize objects by looking at them. However, it is also true that we can touch objects in order to acquire a whole series of information (shape, size, weight and more). Starting from this statement, researchers began to focus on haptics too. As pointed out in Lacey and Sathian's review, two events were important for this direction change: from the 1980s, haptics started to be a stand-alone research field; moreover, from the 1990s, researchers became more and more interested in studying multisensory interactions instead of focusing on separated sensory modalities (Lacey and Sathian, 2014).

On one side, the two sensory modalities seem to present many differences. First of all, haptic processing is restricted to those objects which are inside the peripersonal space, while vision allows to collect information also about those objects that are outside from the peripersonal and quite distant from the observer. Moreover, a big difference is that vision allows us to have an holistic perception of the scene, while haptics requires a sequential exploration. This means that using vision we can explore different objects together and receive more information about the same object in a short time; while using haptics, we will receive a restricted part of the information in a certain time and consequently more time is needed for a big amount of features data. Finally, categorization processes seem to be based on different features for the two sensory systems: shape for vision, texture for haptics (Klatzky et al., 2011).

On the other side, despite the abovementioned differences, vision and haptics seem to draw from a shared object representation. So, unimodal different representations could converge in a common crossmodal object representation.

But, how does this process happen? Is a visual processing necessary in haptic object recognition? Are there other involved mechanisms?

Then, lots of questions remain still open. Moreover, very recently a new research field is conquering the scene, that is about the role played by the gaze (Laeng et al., 2014). In fact, several authors began to wonder if, since visual processing is so important in forming visual, and probably also haptic representations, gaze could have an important part in this process.

The general target of this study is to study the role of visual processing in haptic object perception and representation. Moreover, the idea is to propose an answer for the diatribe about amodal representation/supramodal representation.

For this reason, I will first analyze the available literature on visual, haptic and visuo-haptic object representation and recognition; then, I will present the three experiments I conducted during my Ph.D. studies.

2. Visual object representation and recognition

2.1. Models for visual representation

During the years, many researchers tried to interpret how we succeed in easily and quickly recognize objects visually, despite the big amount of images we are exposed to at the same time, and the fact that these images are constantly changing in several features.

Cooper et al. (1992) suggested to group the various theories about object representation into two categories: "metrically specific" and "metrically invariant" (Cooper et al., 1992). Theories categorized as "metrically specific" are those sustaining that we have an object representation characterized by specific features (size, position, orientation). Thus, when we face an object owning other specific features, a transformation is necessary in order to make these features consistent. This processes will allow the object representation activation. In this category it is possible to find the SCERPO model proposed by Lowe (1985) and the Feature Alignment Model by Ullman (1989). The SCERPO model is based on three different mechanisms, from which it takes its acronym: Spatial Correspondence, Evidential Reasoning and Perceptual Organization (Lowe, 1985). The Feature Alignment Model, instead, foresee two phases: in the first one, there is the creation of an "alignment" between the visual input and existing object models, then, the model that best fit with this alignment is chosen (Ullman, 1989).

Experimental data supporting metrically specific theories are those which showed that participants have worse performance and longer reaction times when they are asked to make a same/different shape recognition tasks with objects varying in size (Bundesen et al., 1981; Jolicoeur & Besner, 1987) or in depth orientation (Tarr & Pinker, 1989; Bulthoff et al., 1991).

In the same direction, more recently, there are studies supporting the existence of a viewpoint-dependence in vision (see Peissig & Tarr, 2007 for a review). Humphrey

& Khan (1992), for example, noticed that participants were able to better recognized 3D novel objects when the test angle was the same as the study angle. Similar results were obtained by Newell and Findlay (1997) with familiar objects too.

Theories categorized as "metrically invariant" support the idea that object recognition does not need a transformation to activate the existing representation, that is independent from the object's specific features. In this category we find the Biederman's Recognition by Components (RBC), according to which objects representation is based on the analysis of object's determinants and their connection in the space (Biederman, 1987). In particular, Biederman proposed a multistage model to identify an object. The first one coincides with the extraction of the object's edges, then it is segmented into diverse concave regions. Now, it is possible to identify the object's components, which will correspond to "geometric primitives" (called "geon" by the author, from "geometrical ions"). A composition of geons and the relation among them activate an object representation; this, in the last stage, leads to the object recognition. Geons are thought to be very simple geometric volumes (like spheres, cylinders, blocks) which are easily recognizable and almost independent from orientation variation and deterioration. Big importance has given to the type of relation among these components, since same geons can identify different objects on the basis of the relation existing among them (Biederman, 1987).

Although the RBC model have not received a confirmation in real life object representation, some studies seem to support it. In object-naming tasks, participants succeeded in quickly recognizing objects which kept just few components, while eliminating the object's outline including geons drove to the impossibility of recognizing the object (Biederman, 1987). Moreover, Biederman and Ju (1988) found similar performances in naming objects presented as an high-definition photograph or a line drawing preserving only some components (Biederman & Ju, 1988).

Hummel and Biederman (1992) proposed the "neural net implementation of RBC" (Hummel & Biederman, 1992). This model forecasts 7 levels: in the first one, cells react to the object image's edges. Then, in the second level, there are groups of cells identifying vertices, axis and blobs. At level 3, geons' features detection takes place.

This information are then used in the following two levels to understand the relation among geons. In the sixth level, geon features assemblies (GFAs) organize in patterns consisting of geons and their relation from the previous levels. Finally, in the last level, the object is recognized.

A particularly interesting model, which tries to account for visual object recognition considering anatomical and physiological limitation, is the one proposed by Riesenhuber and Poggio (1999). This hierarchical feedforward model starts from the "simple-to-complex cells model" described by Hubel and Wiesel (1962) and expands it. In particular, the transition to more complex cells to conquer invariance and specificity proceed trough two mechanisms: "SUM", consisting in the sum of afferents having similar weights (linear mechanism), and "MAX", according to which response depends on the most active afferent (nonlinear mechanism). The first mechanism seems to bring to the features specificity. The latter mechanism, instead, seem to be more efficient in achieving both position and size invariance. This is what Riesenhuber and Poggio hypothesize to happen in their object recognition model. The merit of this new proposed model, then, is to add a MAX mechanism to the known SUM one, and the fact that this nonlinear mechanism seems to receive some support from different studies (Wang et al., 1998; Riesenhuber & Poggio, 1999) as well as being in accordance with neurophysiological data (Sato, 1989) with monkeys.

According to the authors, this model can explain real-situation object recognition, since MAX mechanisms selectively respond to the target object without being distracted by other objects normally forming the scene in real life situations, and that, to a certain extent, are encoded as they are in the receptive fields of the simple cells (Riesenhuber & Poggio, 1999).

As can be seen from the theoretical models exposed up to here, although we easily and quickly recognize everyday visual objects, almost without paying a particular attention or being so concentrated, it is not that simple to understand how this cognitive process takes place. That's why different theories have been suggested trying to shed light on this question.

At the same time, it would be very interesting to know what cortical areas are involved in this complex cognitive process.

2.2. Brain areas involved in visual object recognition

After having described some theoretical models that tried to explain how the complex object recognition process takes place apparently effortless, it is very important to investigate what confirmations there are at a cortical level, and whether these discoveries can help us in better understanding how this process happens in a so effective way.

Over the years, there have been several studies that sought to collect as many information as possible about visual processing. This is a broad concept that can be analyzed by focusing on different aspects: processing of mostly perceptual aspects (for example, Engel et al., 1997, for colors; Tootel et al, 1998, for size), visual attention (see Wojciulik & Kanwisher, 2000 for a review) and object recognition (see Mazer & Gallant, 2000 for a review).

The latter research field is the one I'm aiming to focus on.

Many years have passed since Lashley (1948) proposed that visual processing was confined into the striate cortex, starting from the observation of effects caused by lesions in monkeys' pre-striate areas. To date, it is now clear that this is not true. Rather, various cortical regions are employed in several visual processes. In particular, it is possible to distinguish two hierarchical visual processing streams: the ossipitotemporal/ventral and occipitoparietal/dorsal pathways. The first one (ventral) is also known as the 'what' stream, while the second one (dorsal) as the 'where' stream (Ungerleider & Mishkin, 1982). This because they are supposed to be involved in objects recognition and localization respectively. These findings derive from studies on monkeys' cortical lesions (Ungerleider & Mishkin, 1982; Mishkin et al., 1983) but they also received a reply from human clinical patients (Newcombe et al., 1987). Moreover, thanks to the advent of new research technologies it has been possible to confirm the existence of the ventral and dorsal pathways in humans with studies using the PET technique to measure the regional cerebral blood flow (Haxby et al., 1991; 1993; Kohler et al., 1995; see Ungerleider & Haxby, 1994 for a review).

In this ventral way, studies with monkeys showed that the object recognition process follows a hierarchical pathway which goes from the primary visual cortex (V1) to the inferotemporal (IT) cortex (Tanaka, 1993). With the help of the fMRI technique, some authors (Sereno et al., 1995; Tootel et al, 1996) found that in humans there is a retinotopical organization around the V1, which could be compared to that showed in monkeys V2, V3 and V4. The point is to investigate the existence in humans of an homologous brain area like the IT (inferotemporal cortex) in monkeys, which seems to be fundamental in object representation and recognition.

Malach et al. (1995) proposed that this role is accomplished in humans by the lateral occipital complex (LOC), a region situated in the lateral portion of the occipital cortex and the post-medial area of the fusiform gyrus. Authors compared the cerebral activation, using fMRI, during the passive view of object and texture images. Results showed a greater LOC activation for object pictures than texture pictures, suggesting a primary role of this cortical region in object identification. The following year, Kanwisher et al. (1996) obtained similar results employing line drawings as input stimuli. In fact, is was noted a stronger LOC activation for line drawings representing an object form than for muddled ones. It is important to underline that these results are achieved with both familiar and unfamiliar objects.

Nevertheless the "simple" activation of a selected cortical region during objects view is not enough to assert with certainty that it is specialized for object recognition, since other processes could intervene.

Data supporting the centrality of the lateral occipital complex in object recognition derive from different studies. Some of them concentrated on the investigation about those characteristics linked to object recognition defined as "form-cue invariance" and "perceptual constancy". The first refers to the capacity to identify an object shape independently from the visual cues forming it. The second one, perceptual constancy, concerns the possibility to recognize the object independently from changes in its size or location. Tanaka (1996) demonstrated that activation in macaque inferotemporal cortex (IT) is position/size and form-cue independent. In humans, Grill-Spector and colleagues (1998) comparing the effect of motion, texture and luminance contrast as visual cues, noted that all of them were

comparably able to lead to object identification. fMRI data showed overlapping activation patterns related to the three form-cues, thus supporting the LOC importance in object recognition. Kourtzi and Kanwisher (2000) obtained similar results measuring the BOLD signal for different image cues: contours, shading and monocular depth. Also in this case the LOC proved to be fundamental in object processing, independently from the different eliciting cues.

As to perceptual constancy, Malach et al. (1995) found that changing images' size do not bring to alteration in the cortical activation (LOC), rather there are overlapping activation maps for different stimulus sizes. Interestingly, Grill-Spector et al. (1999), using an alternative technique (fMRI adaptation), advanced the idea that it is possible to identify two subregions in the LOC: an anterior-ventral area, which shows a stronger adaptation together with position and size invariance than the caudal-dorsal area.

Summarizing, the lateral occipital complex seems to have a primary role in object identification and recognition, independently of the visual cues defining objects and of changes in objects' size and location.

Further findings supporting the fundamental role of LOC in object recognition arrive from studies regarding patients with lesions. Damasio et al. (1990), for example, showed that focal lesions in the fusiform region were associated to prosopagnosia, that is the incapability to recognize previous familiar faces and to memorize new ones. Moscovitch et al. (1999), instead reported the case of a man with a LOC injury who was able to recognize faces but he showed an object agnosia, namely he could no longer succeed in object recognition. In a very interesting research, James et al. (2003) took into consideration a patient showing visual form agnosia, but who was at the same time able to process some object information in order to drive actions towards it. This motor capacity versus object discrimination inability resumes the above studied contraposition between "what-ventral" and "where-dorsal" pathways. Moreover, using fMRI technique, authors showed that, contrary to what have been previously found (Malach et al., 1995; Kanwisher et al., 1996), in the patient's LOC there was no greater activation for line drawings representing clear objects than for

muddled line drawings, underlining the object recognition impairment. Finally, high-resolution anatomical RMI confirmed that lesions had hit the LOC area bilaterally.

Summing up, as to cortical regions involved in object representation and recognition, the LOC seems to occupy a prominent place. Different orders of studies, in fact, demonstrated that it plays a fundamental role in the global object form recognition, regardless of visual cues (form-cue invariance) and of changes in input size and location (perceptual constancy). Moreover, LOC impairments (clinical patients with cortical lesions) lead to various types of object agnosia, even while object perception per se remained unaffected.

Thus, even if the processing is not entirely clear, there is the shared idea about the fundamental role played by the LOC in object recognition.

2.3 The role of eyes' movement

In the "Scanpath Theory", Noton and Stark (1971a) referred to the set of saccades and fixations that a person makes while observing a scene using the term "scanpath". Authors suggested, for the first time, that the fixations patterns made while encoding an image and then recognizing it, are almost stackable (Noton & Stark, 1971a).

This statement found a confirmation in studies comparing scanpaths made by participants exposed to more views of the same picture. For example, in a recognition study, Foulsham and Underwood (2008) let participants explore a set of 45 images first. Afterwards, participants received a new group of 90 images and they were asked to say whether they have already previously explored them. Authors noted that scanpaths made during the recognition phase were more similar to the ones made during the learning phase for the same objects, compared to scanpaths for new pictures (Foulsham & Underwood, 2008).

Nevertheless, the Scanpath Theory has often been criticized since eyes' movement patterns were thought to be connected to the visual features of the scanned image (bottom-up control) which capture them, rather than starting from an internal stored sequence.

To prevent this problem, researchers began presenting participants to their experiments with tasks requiring imagery (Brand & Stark, 1997; Holsanova et al.,

1998) and blank spaces observation (Ferreira et al, 2008). Brand and Stark (1997), for example, found almost overlapping scanpaths produced during the exploration and the subsequent imagination of a chessboard diagram. Similar results have been obtained by Holsanova and colleagues (1998) with real life scenes. These results agree with the idea that the repeated scanpaths for the same image are not due to the stimulus presence itself, but rather to internal perceptual patterns. Further data supporting this statement arrives from *imagery* experiments performed in completely darkness (Johansson et al., 2006).

To add further support to the importance of eye movements for scene encoding and recalling, Laeng and Teodorescu (2002), with a particular *imagery* task, showed that when participants were required to keep their eyes fixed on the center while encoding a scene, they tended to do the same during the *imagery* phase. Moreover, as a further confirmation, when participants were instructed to freely explore the scene during the encoding phase and then to fix the center during *imagery*, they showed an impaired performance. According to the authors, this could happen because it is as if a piece of information acquired during encoding is lacking for the mental image reconstruction (Laeng & Teodorescu, 2002).

More recently, Laeng et al. (2014), relying on results from previous researches, assumed that when trying to recover a scene's image previously experienced, not only eye movements take place, but these essay to reproduce the content of the initial scene even if it is not present anymore. In particular, to give evidence of what have been said, in a first experiment they asked participants to visually study an image (a triangle) and then to create a mental image of it while looking at a blank display. In another version of the same experiment, triangles were replaced by animals' pictures. In both cases, results showed that during the mental image construction participants tended to replicate similar scanpaths as in the studying phase and, moreover, the fixation duration was greater in the display areas which contained peculiar features for pictures recognition. In a second experiment, participants were asked to memorize some images (animal pictures) as they would have to respond to same questions about them later. Results showed that when trying to recover the information to provide the answer, participants tended to look exactly at the display region in which the investigated aspect was during the memorization phase. In a

different experimental condition, participants were requested to fixate a point in the display away from the investigated region while finding the answer to the question. This produced worse performances than the previous experimental condition (in which participants could freely move their eyes).

Whether, from one side, it could appear controversial that eye movements occur even when there is nothing to look at or there is nothing to be perceived, from the other they could have an important role. As proposed by Hebb (1968), moving eyes during recalling an information helps rebuilding the scene representation since during perception we also register the eye movements made to acquire the scene. About this, some authors started to sustain that imagery and perception share common mechanisms (Kan et al., 2003) and to join the idea of an "embodied" perception (Barsalou, 1999; Gibbs, 2006), that conceive the perception process as active. In this sense, then, also the imagery process could be considered as active, taking into account the body information in the recovery of internalized data. Above all with visual imagery, it would seem that positions and eye movements could be stored as integral part of the general scene representation (Noton & Stark, 1971a; Noton & Stark, 1971b).

These theories received a good quantity of confirmation from different studies: recognition tasks, imagery tasks in different experimental conditions and neuroimaging studies.

As seen before, in fact, participants' performance was impaired when they were asked to keep their eyes fixed on a specific point or to move them in a different manner from the expected one (Laeng & Teodorescu, 2002; Laeng et al., 2014).

Moreover, Thomas and Lleras (2007) found that "steering" participants' eye movements in a way that suits with the preferable one for the task solution enabled them to provide the right answer, while worse performances were achieved when their eye movements followed a casual pattern (Thomas & Lleras, 2007).

As to neuroimaging studies, they also seem to support this belief, showing that at cortical level perception and imagery seem to elicit similar patterns of activation in the visual cortex. Considering this, Ganis et al. (2004) proposed that this comparable activation could be reflected in the fact that imagining something could claim the

same cognitive processes as when directly perceiving it. In particular, forming a mental image would lead to a scanpath retracing the image itself, even if it is not present in front of the imagining person.

Thus, eye movements field study in connection with scene perception and representation is very intriguing and quite recent, so it needs to be fully explored. Moreover, given the importance they have in visual recognition and representation even in input absence, it would be very interesting to explore whether there could be some interaction with haptic/crossmodal representations.

3. Haptic object representation and recognition

Haptics can be considered a relatively recent research field if compared with vision. The latter, in fact, has always received more attention since it was thought to be the main sensory modality in object recognition, while haptic exploration was less effective and provided only additional information about objects, not necessary for their identification.

Now we know that this is not the case. In comparable conditions, in fact, we are able to haptically recognize objects as effectively as when using vision. Klatzky et al. (1985), for example, showed that participants succeeded in accurately recognize one hundred familiar objects within 3 seconds by using their hands only.

Thus, it is very interesting to explore how this process takes place in order to lead to object representation and recognition.

Before continuing, it is important to clarify that with the word "haptic" we refer to a set of information deriving from object exploration which includes tactile cutaneous perception and kinesthetic data (Loomis and Lederman, 1986).

3.1 Exploratory procedures for haptic object identification

Lederman and Klatzky (1987), in a research destined to stay in haptic exploration history, observed and described those which were then defined by them as "exploratory procedures" (EPs). Authors began wonder if when we use our hands to extrapolate different information about an object, we always make the same movement patterns or if these vary according to the information of interest. Using a match to sample task, Lederman and Klatzky (1987) asked blindfolded participants to choose among three objects the one which best paired the target object according

to a specific selected feature (for example texture, hardness, volume, weight). Performances were videotaped and then scanned frame by frame. They noticed that participants made similar hand movements for the same to be explored object feature. So, when participants freely explored objects, each target feature seemed to activate a specific pattern of hand movements: the exploratory procedures.

Exploratory procedures (EPs) have been defined as follow:

- *"Lateral Motion"*, short and quite fast rubbing movements from left to right and back, usually used to explore texture characteristics;
- *"Pressure"*, a pressing or twisting movement to check if the surface modifies or does not react, usually associated with hardness exploration;
- *"Static Contact"*, passively touch a surface, useful for temperature information;
- *"Unsupported Holding"*, sustaining an object on the open hand and moving it up and down, usually implemented when trying to obtain information about the object weight;
- *"Enclosure"*, trying to surround the object with hands in order to extrapolate information about its volume and general shape;
- *"Contour Following"*, following the contour and borders of the object with the fingers or the hand, usually associated with the specific shape exploration

For a graphical representation of the exploratory procedures described by Lederman and Klatzky see Lederman & Klatzky (1987).

According to Klatzky and Lederman (1993), these exploratory procedures are directly involved in object recognition. This process, in fact, would consist in a series of stages: at each stage one or more reconcilable exploratory procedures are selected and performed in order to obtain knowledge about multiple object features. The result of this process is an object representation construction which is then compared to other representations that are already stored till the best match is found and, consequently, the object recognition occurred.

Lederman and Klatzky (1990) specified that exploratory procedures selection depends also on the assigned task and its aim. Authors observed that the process activated when people were required to determine if an object appertained to a

specific category consisted in two phases: in the first one, participants chose "enclosure" and "unsupported holding" to explore the object feature, probably because these two exploratory procedures provide various general information about the object itself. Then, in the second phase, other EPs were selected according to the specific category aspects to be analyzed (Lederman & Klatzky, 1990).

3.2 Studies with sighted and blind people

As I have previously written, we mainly explore our world by the sense of sight. When, for different reasons (in the dark, blindness), it is not possible to use vision, we should rely on other sensory modalities, like audition and haptics. In object recognition, in particular, the latter seems to be the best candidate (Hatwell, 2003). So, several authors began thinking that to better understand how haptic object recognition works regardless of sight, it could be useful to involve blind people in their studies. Nevertheless, it is not as simple as it sounds. In fact, it is practically impossible to have a homogeneous sample of blind participants, since there are many differences in the modes and periods of their life in which blindness was acquired (Heller, 2000). A common distinction, anyway, is among "congenitally blind", "early blind" and "late blind".

In this subchapter, comparisons among sighted and blind people are done, considering the exploratory procedures as previously described (Lederman & Klatzy, 1987). Moreover, it is important to take into consideration that Lederman and Klatzky (1993) found that when participants were not asked to attend on a specific object feature, there would be a primacy for texture when they used haptics only and, on the contrary, geometric qualities were privileged when they could see the object (Lederman & Klatzky, 1993).

As suggested by Withagen et al. (2012), we could expect blind people to be more efficient in haptic object recognition than sighted ones because of their experience and training in this field, but, at the same time, sighted people could be facilitated in object recognition tasks by the visual imagery support. Moreover, several authors (Sadato et al., 2002; Sathian, 2005; Sathian & Stilla, 2010) showed that the occipital

cortex, usually activated by visual stimuli, is implicated in haptic exploration by blind people too, suggesting brain plasticity (I will treat this topic later).

So, what results could be expected from sighted/blind comparison? Are there differences between young and adults?

3.2.1 Development aspects in exploratory procedures: comparison between young and adult explorers

Almost all exploratory procedures studies involved sighted adult participants. So some questions arise: Are these procedures present during infancy and childhood? When they develop? Are young and adult performances comparable in terms of accuracy and speed?

As to sighted children, Bushnell and Boudreau (1991) observed how infants acquire the necessary skills to use their hands for exploring objects, in order to extract information about their features (size, shape, texture and so on). According to the authors, this process takes place during the first year. In particular, they found that four months old children showed movement patterns which can be associated to the "enclosure" EP defined by Lederman and Klatzky (1987). In the next months (4-9), then, "lateral motion" and "pressure" appear. Nevertheless, those movement patters, albeit similar to those implemented by adults with the known exploratory procedures, seem to have different purposes in children. More recently, Alexander et al. (2002) found an age effect in children facing a comparison (same/different) task. Participants (4-9 years old children) were asked to explore pairs of animals' models (dinosaurs, sharks and whales) and to decide if they were exactly the same or not. Results showed that older children performed better than the younger ones, regardless of their level of knowledge about used stimuli. This was probably due to a more efficient haptic exploration strategy used by older children which led to the conquer of a bigger amount of details useful for comparing models and taking a decision on their similarity (Alexander et al., 2002).

As to blind children, Morrongiello et al. (1994) studied how they and sighted children (3-8 years old) explore and recognize objects. Children were presented with objects changing in size in a haptic recognition task. Results revealed that older children were more accurate and fast than the younger ones, regardless of whether

they were blind or sighted. Authors established older children to be more accurate and precise during exploration, focusing on specific areas besides the global object shape. Moreover, they suggested that sighted and blind children do not differ in their performances, concluding that vision is not indispensable for recognizing objects haptically (Morrongiello et al., 1994). Simpkins (1979) obtained similar results with sighted, visually impaired and blind children (4-7 years old) engaged in an object shape recognition. Performances (recognition accuracy and accurate hand movement patterns) got better with increasing age (Simpkins, 1979). Millar (1997), in general, supported the idea that blind children demand "convergent and redundant information" to reach the same level of knowledge as sighted children (Millar, 1997).

In a recent study, Withagen et al. (2012) employed a similar procedure like Lederman and Klatzky (1987), but they involved in the experiments congenitally blind participants besides sighted ones. Moreover, half of participants were 7-13 years old, the other half were adults. Experiments concentrated on just four object properties: exact shape, volume, weight and texture. Results showed that adults performed better than young in the matching task for all of the object properties, but there was no difference in response speed. As to sighted/blind comparison, authors noted that the only difference between the two groups was that blind adults were faster than blind adults in the "exact shape condition". Withagen et al. (2012) concluded by stating it is likely that adults use more effective and targeted strategies than young; moreover, they gave two possible explanations for speed differences: it could be possible that sighted participants reconstruct mental image while exploring the object shape and this could cause a cost in terms of time, or blind participants hold more relevant strategies to face this task (Withagen et al., 2012).

Summing up, all the previous results are in line with what is explained by Hatwell (2003): young people usually utilized general patterns of exploration which are not the best solution for that object property or for the task solution. Then, growing up, they acquire the exploratory procedures that we know to exist in adults (Hatwell, 2003; Lederman & Klatzky, 1987).

3.2.2 The role of mental imagery

Cattaneo et al. (2008) proposed an exhaustive review about this topic which I suggest to read and from which I will start for my discussion.

As we have just seen, it has not been understood yet what role visual experience plays in haptic object recognition. In fact, we could think blind people to perform better in this task since they are used to rely on their hands to explore objects, so they are more experienced. On the other side, we could reason that sighted people are facilitated in object recognition tasks because they can rely on visual imagery, since they experienced visual processing. The connection between perception and imagery in vision is supported by several behavioral (Kosslyn, 1994; Kosslyn et al., 1999a; Craver-Lemley & Reeves, 1992) and neuroimaging (Kosslyn et al., 1995; Chen et al., 1998, Kreiman et al., 2000; Kaski, 2002) studies, even though the same authors suggested that comparing direct visual perception to visual imagery leads to not so comparable results (Craver-Lemley & Reeves, 1992; Kosslyn et al., 1999b;) or cortical activation level (O'Craven & Kanwisher, 2000). Chen and colleagues (1998), for example, tested sighted participants with two mental imagery tasks ("hometown walking" and "flashing light") in the dark. fMRI data showed that both visual perception and imagery activated the visual cortical areas; in particular, V1 (primary visual cortex) and the lateral geniculate nucleus were the regions involved. These results supported a shared brain activity for perception and imagery with visual stimuli. Kosslyn et al. (1999b) solicited participants first to memorize a set of stripes and then to compare a pair of them about a requested feature (like extent) by using imagery only or by seeing the stimulus image. During the test phase, two different techniques were used: PET and rTMS (repetitive transcranical magnetic stimulation) obtaining complementary results. Authors concluded, in fact, that PET showed the medial occipital cortex (in particular area 17) activation during both test conditions (visual imagery and visual perception) and rTMS data confirmed this statement by showing that when applied to that specific brain area this technique resulted in a performance decline in both test conditions (Kosslyn et al., 1999b). Similar results were obtained by Kreiman et al. (2000) when epileptics patients were asked to visualize a previously seen image. Results led authors to support a shared neural substrate for visual perception and mental imagery (Kreiman et al., 2000).

Kaski (2002), in a useful review, suggested that visual perception is necessary for visual imagery to take place. These results were supported also by Ganis et al. (2004) who, from one side, found that answering to questions about a mentally visualized or perceptually seen object picture activated a common neural substrate.

Thus, considering the studies above reviewed, should we consider blind people to be irreparably disadvantaged?

Fortunately, this does not seem to be the case, advancing the thought that they (and mental imagery in general) could lean on other mechanisms than visual processing. Several studies, in fact, supported that different brain areas are engaged by visual perception and mental imagery. Ganis et al. (2004), about that, noticed that although there was a comparable general activation, this was not homogeneous, with greater overlapping in the parietal and frontal areas rather than in occipito-temporal one. This would suggest that higher-order mechanisms could work with visual perception and mental imagery in a similar manner, while lower-order mechanisms could differ between them (Ganis et al., 2004). Relying on ERPs (event-related potentials) data, Farah (1989) suggested that visual imagery seems to engage the visual association cortex instead of the primary visual cortex. This result was confirmed by D'Esposito et al. (1997) with the help of fMRI, a more efficient technique with this type of information. Authors asked participant to listen to an object's name and visualize it (condition 1) or just wait for the next one (condition 2). Comparing the neural activation during the two experimental conditions, D'Esposito et al. (1997) confirmed what previously suggested by Farah (1989): mental imagery do not directly activate the primary visual cortex, but it seems to involve the visual association areas (in particular area 37) (D'Esposito et al, 1997). Another confirmation derives from the similar results get by Knauff et al. (2000): the authors, again, support the idea of a mental imagery relying on higher-order visual brain regions, like the visual association cortex, while the primary visual cortex is not necessarily implicated (Knauff et al., 2000).

Further support to these findings comes from studies involving patients with cortical lesions. Behrman et al. (1992), opposing to the existence of a shared representation for visual perception and imagery, described a patient showing agnosia, namely the

difficulty to recognize visually explored objects. Interestingly, mental imagery was preserved: he was able to solve various tasks requiring the mental creation of an object image. These observations led the authors to suggest that mental imagery rely on a different more complex representation than the visual perception per se (Behrman et al., 1992). In the same direction, Bartolomeo et al. (1998) reported that a patient with damaged temporo-occipital areas had a series of difficulties linked to visual perception (agnosia, alexia, achromatopsia) but an unharmed visual imagery. Thus, according to the authors, mental imagery and visual perception seem to recruit different brain regions (Bartolomeo et al., 1998).

Important information come from the study by Chatterjee and Southwood (1995) with three cortical blind participants. Also in this case, it was demonstrated that patients were able to solve imagery tasks and the primary visual cortex was not implied, suggesting that visual experience and mediation is not necessary for the visual imagery to work and that they rely on different mechanisms and brain areas (Chatterjee & Southwood, 1995).

Aleman et al. (2001) made a comparison between sighted and congenitally blind participants in a mental shape comparison task. Authors asked participants to compare three spoken shapes and to choose the one to be excluded (namely the one which was more different from the others). Results showed that, even though congenitally blind made more mistakes than sighted, the first properly succeeded in solving the task. Aleman and collaborators (2001) suggested that blind participants could rely on haptic imagery as well as sighted ones rely on visual experience, also concluding that a visuo-haptic shared representation exists. Noordzij et al. (2007), using a very similar method like Aleman et al. (2001), compared for the first time in the same task, sighted, early-blind and late-blind participants. They found no differences between sighted and early-blind participants' performances and, consequently, suggested that blind people rely on haptic imagery and this goes in the more and more shared idea that vision loss is compensated with the experience in other sensory systems (Noordzij et al., 2007).

Further support also derives from studies indicating that brain regions always considered to be predominantly visual, can respond to haptically presented stimuli.

Pietrini et al. (2004), for example, found that recognizing common objects visually (sighted) or haptically (blind) engaged the same cortical areas, namely the ventral visual pathway. According to the authors, this result cannot be linked to the employment of visual imagery in haptic recognition, but rather to the existence of a sensory-independent representation located in the visual cortex (Pietrini et al., 2004).

Another suggestion to explain the engagement of visual cortical areas during haptic processing comes from studies suggesting a functional cortical reorganization in blind people (D'Angiulli & Waraich, 2002), a topic that I will treat in the following subchapter.

It is very interesting to say something about dreams' content in blind people. Bertolo et al. (2003) and Bertolo (2005) compared sighted and blind people on dreams' content drawings and alpha rhythm data. They found no differences between the two participants formations: blind were able to accurately draw the previous verbally described dream's scene. Thus, according to the authors it is possible to state that congenitally blind present visual material in their dreams, even if they did not have any previous visual experience, and this suggests a sensory invariant visual imagery (Bertolo, 2005).

3.2.3 A functional cortical reorganization

Several studies showed that an early visual loss can be followed by a functional cortical reorganization in which the one representing the "visual" cortex in sighted people, can be employed in processing stimuli coming from different sensory systems than vision (Uhl et al., 1994; Amedi et al., 2003; Collignon et al., 2007).

As to haptic/tactile stimuli, which are in my interest, different studies with the application of different technologies seem to get to the same result: typically somatosensory inputs activate visual brain regions in blind people only. Sadato et al. (1996), for example, measured how much visual cortex activates in early-blind participants compared to sighted, during a Braille discrimination task using PET technology. They showed that the occipital regions activated during the task only in blind participants. Cohen et al. (1997), instead, used the transcranical magnetic stimulation (TMS) in association to haptic identification tasks with sighted and early-blind participants. Authors showed that applying the TMS to the occipital cortex

(implied in visual processes) affected blind participants' performance but not the sighted one, who (according to the literature) would have been affected when facing visual tasks (Amassian et al., 1998). Cohen and collaborators, then, suggested a cortical reorganization in blind people, with the implication of visual brain regions in haptic processing (Cohen et al., 1997). Amedi et al. (2005) obtained similar results using fMRI: they found visual cortex activation in relation to Braille reading in congenitally blind participants compared to normal sighted people.

At this point, a question arises spontaneously. Does brain plasticity occur during a specific developmental period of life? Some studies demonstrated that there seems to be a time frame in which the brain reorganization can take place (Cohen et al., 1999; Sadato et al., 2002; Sathian, 2005). Cohen et al. (1999), for example, compared congenitally blind, early-blind and late-blind participants in a Braille reading task, with the utilization of two techniques: PET and rTMS. PET data showed no activation in the occipital cortex for late-blind participants only. In the same way, rTMS applied to the same areas caused impairments in Braille reading in the other participants (congenitally- and early-blind). In the authors' opinion, thus, there is a period in which this brain reorganization seems to be possible, and it stays within 12-14 years (Cohen et al., 1999). Sadato et al. (2002), using fMRI technique, showed that during passive tactile Braille dots discrimination it is possible to find a greater activation in the visual association areas in all of the blind participants than in sighted ones. As to the primary visual cortex, results differ between early-blind and late-blind (the watershed is 16 years): authors found an activation only in the first group. Thus, even if there is no agreement on the specific age, it is a shared idea that brain plasticity presents a critical period in which it can take place. Whether, on one hand, this sounds like bad news for late- blind participants, on the other it could stimulate research to grow in the perspective of creating sensory substitution devices for congenitally and early-blind people.

Summing up, according to Cattaneo et al. (2008), researches which paid their attention to the study of visual areas implied in haptic processing by blind people, using different brain activity investigation techniques, highlighted two possibilities: the existence of supramodal cortical regions, in which inputs are processed regardless of their sensory modality located in the occipito-temporal cortex;

otherwise, the functional cortical reorganization led visual areas to be engaged in haptic processing by blind when visual stimulation is not present.

4. Visuo-haptic object representation and recognition

In the visuo-haptic search field, there are not so many studies investigating the crossmodal object representation and recognition; so the existing literature is shared by different authors. Very exhaustive publications concerning this topic are those by Lacey and Sathian (Lacey & Sathian, 2012; Lacey & Sathian, 2014). Therefore, I will proceed following their logical structure, but analyzing and concentrating on the most relevant researches that constitute the basis of my studies (crossmodal visuo-haptic object representation).

4.1 Visuo-haptic object categorization

First of all, it is important to say something about the categorization process, since it seems to be the basis of good recognition processes. Nevertheless, as we already said in the introduction, almost all studies focused mainly on the visual system, leaving little space for the investigation of haptic and visuo-haptic categorization processes. Very recently, anyway, several authors started to be interested in studying similarities and differences between haptic and visual categorizations. Cooke et al. (2007), for example, concentrated on the importance of similarity in categorization processes and they wondered how this concept works when people are asked to rate abstract 3D items changing in shape and texture by using vision, haptics or both modalities together. Similarity analysis data revealed that, in all of the cases, the "perceptual space" seems to be comparable for the two systems (Cooke et al., 2007). Similar results were obtained by Gaißert and colleagues using complex shell-shaped stimuli (Gaißert et al., 2008) and real seashells (Gaißert & Wallraven, 2012). It is important to underline that the same result was obtained again when Gaißert et al. (2010) compared passive viewing of 2D objects to haptic exploration of 3D ones, so that they had similar investigation modalities for vision and haptics (Gaißert et al., 2010).

However, other results from the same abovementioned authors suggest that the categorization process is not superimposable for vision and haptics, since objects features have different weights in different modalities. In particular, Cooke et al. (2007) found that shape was the dominant feature in visual condition, while it has a weight comparable to texture in haptic and bisensory conditions (Cooke et al., 2007). Gaißert et al. (2010), using more complex objects, noticed that only shape features had different weights for the two sensory modalities.

As proposed by Lacey and Sathian in their review "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" (Lacey & Sathian, 2014, p. 3).

At the same time though, recently Yildirim and Jacobs (2013) found in their experiments that participants were able to shift knowledge related to categorization across vision and haptics (Yildirim & Jacobs, 2013), suggesting that a shared multisensory representation could be hypothesized for crossmodal categorization. Thus, we do not have a final answer about this contraposition and more researches are needed.

4.2 Cross-modal recognition asymmetry

It is now known that people are able to visually recognize objects previously explored haptically and vice versa. But two questions arise: is crossmodal object recognition comparable to unimodal object recognition? Does crossmodal recognition process obtain the same results independently from which sensory modality comes first? According to some authors, the answer seems to be 'no' in both cases. Casey and Newell (2007), for example, using a face recognition task found that performances were better in within-modal recognition condition than in the crossmodal one, suggesting different representations for the two sensory systems (Casey & Newell, 2007). Similar results have been obtained by Bushnell and Baxt (1999) with 5 years old children (Bushnell and Baxt, 1999) showing lower accuracy for crossmodal recognition than the within-modal one. As to the second question, in

crossmodal recognition people gain better results when vision comes first, namely when objects are encoded visually and recognized haptically (Streri & Molina, 1994; Lacey & Campbell, 2006). According to Lacey and Sathian (2014), two reasons could explain this cross-modal asymmetry: first, information about the object shape could be better encoded by vision (as we saw above) than haptics so this would make the visual-haptic condition easier than the opposite condition. In the latter case, in fact, people performance could be affected by the interference of other objects features that seem to be more relevant for the haptic system (like roughness, plasticity). Second, the memory capacity and duration could vary for visually and haptically explored objects. About that, Bliss and Hämäläinen (2005) using a letter recognition and letter memory task demonstrated that tactile working memory appears to have a more confined and inconstant capacity than the visual one in young adults participants (Bliss & Hämäläinen, 2005). At the same time, about this topic, Woods et al. (2004) conducted a very interesting study, suggesting that haptic working memory (like visual working memory) could be more durable in time. In crossmodal object matching tasks, authors found that between the two experimental conditions (visul-hapic and haptic-visual) there was no difference in performances. Moreover, the same result was obtained by comparing performances in unimodal matching tasks, though general accomplishment was greater in the visual situation than in the haptic one (Woods et al., 2004). Congruent results were obtained by Pensky et al. (2008), who showed that visual and haptic information deteriorated basically in a similar way along 7 days. Summing up, it seems that differences in working memory capacity and duration alone are not sufficient to explain this cross-modal asymmetry.

Also neuroimaging studies provide some information about asymmetries between vision and haptics in crossmodal object recognition tasks. In particular, Kassuba et al. (2013) noticed that bilateral object-specific LOC, fusiform gyrus and intraparietal sulcus are more activated during the crossmodal visual-haptic condition than the unimodal one. According to the authors "results indicate a directional asymmetry in crossmodal matching of visual and haptic object features with a functional primacy of vision over haptics in visuo-haptic object recognition" (Kassuba et al., 2013, p. 68).

4.3 The investigation of neural substrates

At this point, it is interesting to understand what happens at the level of neural substrates. In 1995, Malach et al. showed their participants some pictures of objects, textures and interfering stimuli; using the fMRI technique authors found that a specific brain area, located in a region named "lateral occipital complex (LOC)", selectively activated during the exploration of objects pictures (Malach et al., 1995). Using the same methodology, Amedi et al. (2001; 2002) confirmed these results using 3D stimuli and, in addition, they found that preference for objects in this area can be seen for both visual and haptic stimuli (see also Stilla & Sathian, 2008; Zhang et al., 2004). Moreover, they demonstrated an overlapping between most of somatosensory object-related voxel and the visual ventral pathway. Authors titled this area "lateral occipital tactile-visual region (LOtv)". These conclusions seem to substantiate the supposition that these cortical areas could form a multimodal/supramodal object representation, and may thus be implicated in the processing of object shape information independently from the sensorial system used to encode the stimulus (Cattaneo et al., 2008).

Confirmations about the importance of this area in visuo-haptic shape processing come from neurological case studies. Feinberg and colleagues (1986) described a patient with lesion in the occipital and temporal lobes to be no longer able to recognize objects both visually and haptically (visual-haptic agnosia), though he could accomplish other visual and haptic tasks (to copy a figure, unimodal/crossmodal match; Feinberg et al., 1986).

Other brain areas have been studied relatively to multimodal object shape processing. Among these, Stilla and Sathian (2008) pointed out that the areas activated during bimodal (vision and haptics) shape processing were: the postcentral sulcus and parts (ventral, anterior and posterior) of the intraparietal sulcus (IPS). It is important to keep in mind that most of the cited areas were primarily found in non human primates, so the correspondence with the human brain remains dubious and their role in visual-haptic shape processing needs to be accurately explored.

In summary, the exposed results seem to strongly suggest a visual cortical areas recruitment during haptic processing; anyway, it remains unclear if this leans on top-

down processes implying visual imagery or on bottom-up processes which involve modality-invariant representations.

4.4 Does a shared representation exist for vision and haptics?

To understand whether vision and haptics share a common shape representation, it is important to investigate what role visual imagery could potentially play and if a modality-independent shared representation is plausible.

4.4.1 What's the role played by visual imagery?

Reading the results of the above presented researches a question may arise. Why does haptic perception evoke activation of visual cortex? The easiest answer seems to be, as suggested by Sathian et al. (1997), that haptic processing leans on visual imagery. This hypothesis is sustained by studies noticing that mental imagery activates the lateral occipital complex (LOC) and visual associative areas. In this sense, De Volder et al. (2001), using the PET technique, showed that imagining objects shape leads to responses in occipito-temporal and visual association areas (in particular the Broadmann area) in both early-blind and sighted participants (De Volder et al., 2001). Newman et al. (2005) obtained similar results using the fMRI techniques during a mental evaluation task concerning material and geometrics properties of objects (Newman et al., 2005). Finally, participants in Zhang et al. (2004) experiments reported common experience of visual imagery during haptic shape perception (Zhang et al., 2004).

Nevertheless, Amedi et al. (2001), noticing a higher activation in occipito-temporal regions when people directly explore the object compared to imagery, proposed that visual imagery by itself cannot account for "visual" brain areas recruitment during haptic perception (Amedi et al., 2001). In the same direction, we find studies comparing sighted, early-blind and late blind participants. For example, Pietrini et al. (2004) using fMRI to register brain regions activation during visual (for sighted) and haptic (for blind) recognition tasks, found that "visual" areas were recruited even in

blind people. Stilla et al. (2008) obtained similar results with early-blind and late-blind participants (see Sathian and Lacey, 2007 for a review).

Taken together these results propose that visual imagery alone fails in explaining the engagement of the ventral visual pathway in haptic representation, at least concerning early blind people. In fact, we can't leave out that it could hold an important role in sighted people, mostly if we take into consideration all studies referring to crossmodal brain plasticity following the sight loss (D'Angiulli & Waraich, 2002; Noordzij et al., 2007; Sathian & Lacey, 2007).

As we already discussed in the previous chapter, also, mental imagery is not obligatorily linked to visual processing or visual experience in general (Craver-Lemley & reeves, 1992; O'Craven & Kanwisher, 2000; Aleman et al, 2001) but rather the existence of a shared visuo-haptic higher order mechanism is more plausible (Behrman et al. 1992, Knauff et al., 2000; Ganis et al., 2004).

4.4.2 An amodal shape representation

Another answer to the question made in the previous subchapter, is that vision and haptics effectively rely on a shared representation in which the two modalities input converge. Lacey and Sathian (2012) "suggest the use of the term "multisensory" to refer to a representation that can be encoded and retrieved by multiple sensory systems" (Lacey & Sathian, 2012, p. 3).

This hypothesis found bases on different researches. Peltier et al. (2007) for example, used discrimination visual and haptic tasks in addition to fMRI to show that, with haptic processing, the LOC, the IPS and the postcentral sulcus (PCS) were active; the first two regions responded also in visual condition. Then, authors proposed a model that foresee two streams: a bottom-up one going from the PCS to the IPS, and a top-down path from the IPS and the LOC to the PCS (Peltier et al., 2007). In the same direction, Deshpande et al. (2008) found both a bottom-up path from the left PCS and right posterior insula to the right LOC and the occipital cortex, and a top-down path from the parietal area to the LOC (Deshpande et al., 2008). Lucan et al. (2011) used high-density EEG to calculate the timing of haptic object recognition operations. They again confirmed that the LOC has a multisensory role in visuo-haptic object recognition (Lucan et al., 2011). Pietrini et al. (2004), in the already

mentioned study, found that haptic object recognition activate a "posterior inferior temporal region of visual extrastriate cortex" (Pietrini et al., 2004, p. 5662) and that these response patterns can be compared to the ones elicited by visual recognition with the identical items.

Pulling the sums, activation of the visual regions seems to be more linked to an abstract representation of objects than to the brokerage of a visual mental image.

Other demonstrations for an overlapping of visual and haptic representations come from studies on unimodal and crossmodal priming. Considering behavioral studies, Easton and coll. (Easton et al., 1997a; Easton et al., 1997b) demonstrated the existence of comparable effect for crossmodal and unimodal priming with verbal material, 2D patterns and 3D common objects. Similar results were obtained by Reales and Ballesteros (1999) with picture-fragment completion and object decision tasks. Neuroimaging studies also confirm this findings: James et al. (2002b) noticed that there is not any difference in the efficacy between visual (unimodal) and haptic (crossmodal) priming while viewing at primed and not primed object. In fact, they found a greater activation both for visually and hapically primed objects in visual regions than for not primed objects.

4.4.3 Unimodal viewpoint-dependence vs crossmodal viewpoint-independence

First of all, let's define the terms "viewpoint-dependent" and "viewpoint-independent/viewpoint-invariant". We can say that a representation is viewpoint-dependent when the object recognition is better if it comes from the same angle as the encoding. On the contrary, a representation is said to be viewpoint-independent when the object recognition is equally good regardless of whether the point of view from which it takes place is the same or vary from that of the encoding.

What can we say about visual, haptic and crossmodal object representation?

As to the visual system, several authors showed that visual object representations appear to be viewpoint-dependent (see Peissig & Tarr, 2007 for a review). Tarr and Pinker (1989), for example, taught participants to quickly name some objects from a specific angulation; then, they presented those objects from the same angle as

participants learnt them or from a different one. Results showed that reaction times (RTs) were significantly longer when observation's angles did not coincide. Likewise, Humphrey and Khan (1992) noticed that participants were able to better recognized 3D novel objects when the test angle was the same as the study angle.

As to the haptic system, this topic did not receive the same attention as in vision. Anyway, some authors proposed that, like for vision, also haptic object representations are viewpoint-dependent. In particular, Newell et al. (2001), using small unfamiliar 3D objects (compositions of LEGO bricks), found that participants' performance in haptic object recognition is better if they can explore objects from the same orientation they did during the learning phase. This orientation, according to their results, means the back of the object in the haptic exploration, while for vision it usually corresponds to the front of the object. Authors argued that it is as if haptics allows to explore what remains hidden for vision (Newell et al., 2001). Lacey et al. (2007) replicated these results using a similar methodology.

At this point, it is interesting to note that there is a resemblance conjuring vision and haptics: both of them present "canonical views" (Woods et al., 2008). This means that there are preferential "views" which participants choose for exploring the object and that allow to better succeed in the object recognition. By going to compare these canonical views, anyway, it is possible to observe that they differ in the two sensory systems. In fact, while in vision the best angle is when the median axis of the object is rotated 45 degrees with respect to the observer's body midline (Palmer et al., 1981), for haptic objects we have the "canonical view" when the object's axis and the observer's body midline are parallel or perpendicular (Woods et al., 2008). In the same direction, as we have already seen above, Newell et al. (2001) noticed that in haptic object recognition the preferential "view" is the back of the object while, on the opposite, in vision it is the frontal side of the object.

From the displayed results, then, it would seem that vision and haptics deal with object representations in similar ways, but, at the same time, the processes leading to this could be different for the two sensory modalities.

Therefore, what happens in crossmodal object representation? Interestingly, Lacey et al. (2007), using a recognition task with non-familiar objects, showed that rotating

objects between the study and the test phases impaired participants' performance in the unimodal (visual or haptic) condition, but not in the crossmodal one (visuo-haptic and haptic-visual). In the same year, Ueda and Saiki (2007) obtained similar results, supporting a viewpoint-dependence for unimodal object representation and a viewpoint-independence for crossmodal object representation. Authors, anyway, specified that the viewpoint-dependence/independence for unimodal recognition depends on whether participants were/were not previously informed about the test modality. If they were, unimodal recognition conditions appeared to be viewpoint-dependent (Ueda & Saiki, 2007; 2012); if they were not, as for crossmodal recognition condition, it was viewpoint-independent (Ueda & Saiki, 2007). Although this result may seem strange, Ueda and Saiki (2012) later showed that this could be linked to eye movements' patterns. In fact, when participants knew they were asked to study an object visually and to recognize it haptically (visuo-haptic crossmodal condition), fixations during the study phase were more widespread and protracted than during the unimodal conditions. Lacey and Sathian (2014) hypothesized this explanation could work also for unimodal recognition when participants do not know which sensory modality they are going to be requested to use for the test phase. In this way, participants try to collect as much information as they can about the to be recognized object so that they will be ready to answer in both cases.

At this point, Lacey et al. (2009b) tried to understand how unimodal viewpoint-dependent representations transform in a crossmodal viewpoint-independent representation. Is it a direct process? Otherwise, does it need an intermediate processing? Starting from results obtained in two perceptual learning/object recognition experiments, the authors proposed a model in which unimodal viewpoint-dependent representations straight converge in a crossmodal viewpoint-independent object representation, without the intermediate transformation of unimodal viewpoint-dependent object representations into viewpoint-independent ones. In fact, they noted that perceptual unimodal learning in one sensory modality lead to viewpoint-independence not only in that modality, but also in the other. Moreover, perceptual crossmodal learning reflected in unimodal viewpoint-independence (Lacey et al., 2009b). Lacey and Sathian argued that these results could be explained by the "integration of multiple low-level, view-dependent,

unisensory representations into a higher-order, view-independent, multisensory representation" (Lacey & Sathian, 2012, p. 6).

4.5 A model for visuo-haptic object representation

Lacey and colleagues (2009a) advanced a model for visuo-haptic object shape representation, in which a multimodal representation is located in the LOC. This representation doesn't care about the input sensory modality and can be accessed by two distinct pathways (top-down and bottom-up) depending on whether the processed object is known or novel. The top-down pathway goes from parietal and prefrontal areas to the LOC and it is activated by familiar objects processing, while the bottom-up pathway proceeds from the somatosensory cortex to the LOC and it is implicated by unfamiliar objects processing. Moreover, familiar object processing seems to use the visual imagery, while unfamiliar object processing recruits the spatial imagery (Lacey et al., 2009a).

This proposed model finds its basis on the literature reported in the previous subchapters but, above all, on a series of studies focusing on inter-task correlation of activation magnitudes (Lacey et al., 2010a) and connectivity (Deshpande et al., 2010) related to visual object imagery and haptic shape discrimination tasks. In the imagery task, participants heard to two objects' names and they were asked to say if their shapes were the same or not. In this way, it is sure that participants accomplish the visual imagery task. Then, participants faced an haptic shape discrimination task with known or novel objects. Results showed that object familiarity is the discriminating aspect. In fact, only with known objects authors registered an inter-task correlation of activation for the two tasks and only in this case there was an overlap of cerebral regions activation (connectivity analysis). Specifically, as foretold by the model proposed by Lacey et al. (2009a), two different pathways seemed to be activated: a top-down stream starting from prefrontal and parietal areas directed to the LOC for familiar objects, and a bottom-up stream going from the somatosensory cortex to the LOC for unfamiliar objects (Deshpande et al., 2010).

These data (Deshpande et al., 2010) were recomputed by Lacey et al. (2014) with the help of more updated connectivity technology in a study focusing on the role played by spatial/object visual imagery in relation with object familiarity. Results seemed to be in agreement with the model proposed above (Lacey et al., 2009a). In effect, there was a wider connectivity between object visual imagery and familiar haptic object discrimination on one side, and between spatial visual imagery and unfamiliar haptic object discrimination on the other (Lacey et al., 2014). This is probably due to the fact that novel objects need to be wholly explored in order to allow a representation to be formed (Lacey et al., 2009a). In this process, an important role seems to be played by the IPS (intraparietal sulcus), which is involved in forming an object representation by combining the spatial object information (bottom-up processing; Lacey et al., 2014).

On the other hand, this does not seem to be the case for known/familiar objects, in which a less detailed exploration actives an existing representation via top-down processing starting from the prefrontal cortex (Lacey et al., 2009a).

5. Experiments

The general aim of my research is to understand what role the visual processing plays in haptic object representation. Moreover, I will try to propose an answer for the diatribe about amodal representation/supramodal representation.

In particular, it is true that several studies already supported the existence of a shared representation for vision and haptics, but almost none of them demonstrated its being by using abstract 3D objects, which cannot provide any familiar prompt and, furthermore, none of them brought support for an amodal or supramodal representation.

5.1 Experiment 1 – Crossmodal Recognition

The aim of this first experiment was to investigate if vision has a primary role in recognizing abstract 3D objects and if crossmodal priming has similar/different effects in visual and haptic object representation and recognition.

Based on the abovementioned literature, the hypothesis is that sensory modalities order has different effects on recognition accuracy. In particular, visual priming will lead to a higher haptic object recognition than the opposite condition.

5.1.1 Method

Participants: 84 healthy university students (56 men and 28 women) from Lawrence Technological University (Southfield, Michigan), with a mean age of 20.73 years (range 16-34 years), voluntarily participated in this experiment in exchange of university credit or a small amount of money. All of them read and signed (or their legal guardians in the case they were minors) the informed consent. Participants reported normal or corrected to normal sight and touch.

Materials and apparatus: A total of 12 white plaster sculptures were selected for the experiment from a broader list of abstract items created by students of a basic design class at Lawrence Technological University. The sculptures were selected in

pairs of two by similarity. We selected 4 non-classifiable shapes, 4 cube-like shapes, and 4 sphere-like shapes. They were fixed on wooden stands by means of a vertical stick, both painted black (see Figure 1 for an example). All the objects were unfamiliar, without any explicit meaning and function and hence arduous to be labeled. Six of them, defined *target* sculptures, were presented both visually and haptically during the experiment (Figure 2a) while the other six (defined *distracting* sculptures) were presented either haptically or visually (Figure 2b).

Sculptures were fixed with Velcro® straps to a 242x14x2.5 cm wooden plank on the top of a 304x46x74 cm table.

Figure 1. An example of sculpture

Figure 2a. Target items

Figure 2b. Distracting items

Procedure: The recognition experiment consisted of two between-subjects experimental conditions: the haptic-visual condition and the visual-haptic condition. Half of participants participated to the haptic-visual condition in which the haptic

exploration of the six target items was followed by the visual exploration of twelve items (the six items previously explored haptically plus six new items). The other half of participants took part in the visual-haptic condition in which the visual exploration of the six target items was followed by the haptic exploration of twelve items (the six items previously explored visually plus six new items).

Participants were not explicitly told that it was a recognition task at the beginning.

For the haptic-visual condition, participants were blindfolded and guided in the experimental room in which six sculptures were placed on the wooden plank. They were guided in front of the first of the six items and their hands were positioned in the space immediately above it. They were then invited to slowly lower down their hands until reaching the item. Therefore, they were allowed to freely manipulate the items with two hands and with no restrictions, except picking up the objects. After 20 seconds of exploration participants were asked to raise their hands and to wait for the next object. Participants rested for 2 seconds and then they were guided towards the next item and prompted to manipulate it. This procedure was repeated for all the six sculptures (Fig. 3). After tactile exploration was completed, participants left the room and waited for 2 minutes.
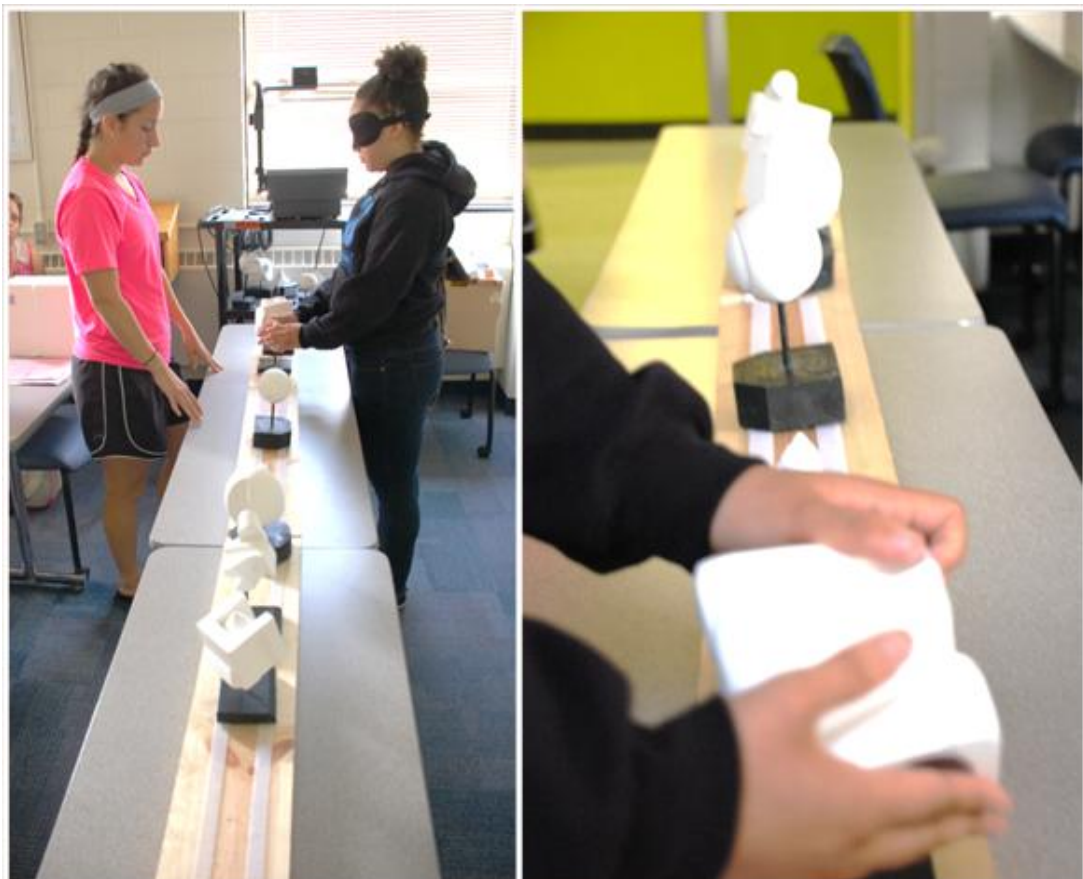

Figure 3. A participant during the haptic exploration of the 6 target items

While the participant waited outside, all the 12 sculptures (the six that were just explored haptically plus six new ones, never presented before) were arranged on the wooden plank in a random order. Sculptures were covered with white plastic boxes to prevent their vision. The order of the twelve sculptures was randomized between participants. The participant was then led back into the room, and placed behind the first sculpture's position on the table. The white box for just that sculpture was removed for 10 seconds, during which time the participant was allowed to visually examine the sculpture but not to touch it. The white box was then placed back over the sculpture, and the participant was asked to answer whether he/she had previously touched the sculpture during the haptic block. Once the answer was provided, the participant was led to the next sculpture's box and the process repeated for each of the remaining 11 sculptures.

For the visual-haptic sequence the sensory modalities were reversed. Participants first explored visually each of the six target items for 10 seconds. Then they were guided outside the experimental room for 2 minutes, the time experimenters needed to add the other six items and randomize all twelve items' positions. Participants wore a blindfold before entering the experimental room for the second time. They explored each item haptically for 20 seconds and asked to say if they saw it in the previous (visual) block.

The two experimental conditions were randomized among participants.

5.1.2 Analysis

At first I ran an ANOVA analysis with *Modality Order* (haptic-visual vs. visual-haptic) as between-subjects factor and *Accuracy* as the dependent variable to test the different level of recognition accuracy for the two sensory modality sequences.

Then, signal detection theory (SDT) was calculated for the accuracy data to evaluate the sensitivity (d') and the presence of a response bias (criterion). To test the sensitivity and response bias of participants in the recognition task, I ran two ANOVA analyses with *Modality Order* (tactile-visual vs. visual-tactile) as between-subjects factor and d' and *Criterion* as the dependent variables.

5.1.3 Results

Accuracy for cross-modal recognition was influenced by the order in which the two sensory modalities were presented. Specifically, the accuracy was 64.6% for haptic-visual order and 72.6% for visual-haptic order. Analyses showed that accuracy for visual-haptic sequence is significantly higher than accuracy for haptic-visual sequence, $F(1, 82) = 5.65$, $p = .020$, $\eta_p^2 = .064$.
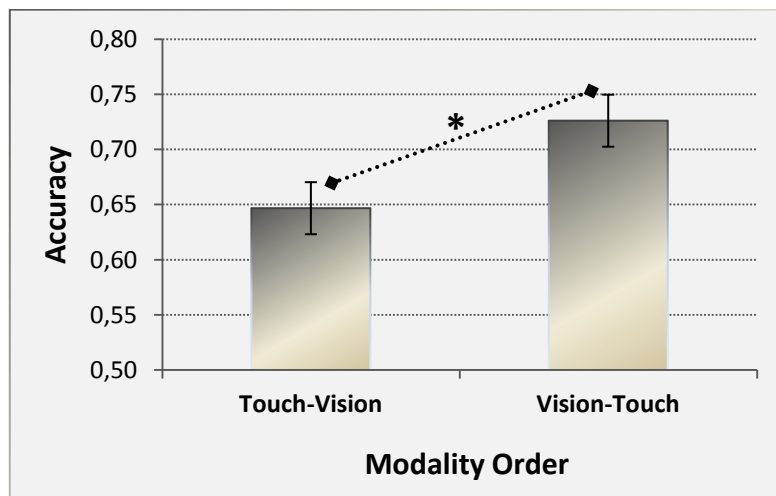
Two single sample t-test were conducted to determine if d' and *Criterion* values were significantly different from zero. Results show that the d' values for the touch-to-vision sequence ($M = 1.37$, $SD = 0.28$) and for the vision-to-touch sequence ($M = 2.08$, $SD = 1.92$) are significantly different from zero, indicating that participants perform above chance level in both sequences [$t(41) = 4.82$, $p < 0.001$ for touch-to-vision sequence and $t(41) = 7.04$, p < .001 for vision-to-touch sequence]. The *Criterion* values for the touch-to-vision sequence ($M = -0.06$, $SD = 0.17$) and for the vision-to-touch sequence ($M = 0.26$ $SD = 0.97$) are not statistically different from zero, indicating that there is no response bias.

Concerning cross-modal recognition sensitivity, results show no significant effects for *Modality Order*, $F(1, 82) = 3.02$, $p = .086$, $\eta_p^2 = .035$ indicating apparently that recognition sensitivity is not influenced by the order in which semsory modalities are presented, even if the p value (.086) suggests a trend in this direction.
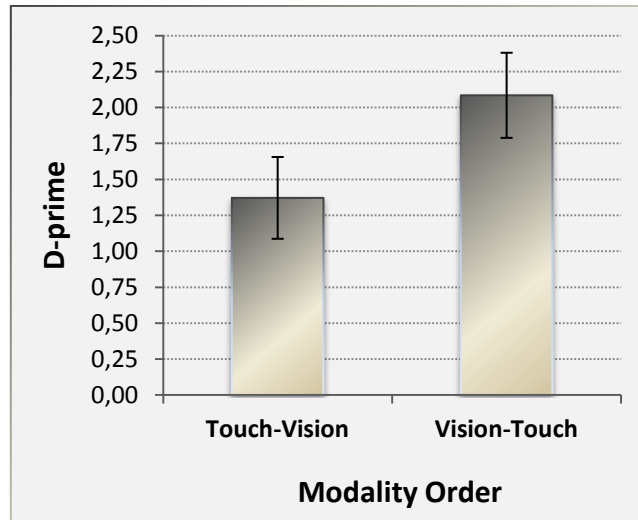
For criterion too, results show no significant effects for *Modality Order*, $F(1, 82) = 2.00$, $p = .161$. This suggests that response approach (liberal vs. conservative) was not influenced by the sensory modality order.

5.1.4 Discussion

The purpose of this study was to investigate if vision has a primary role in recognizing abstract 3D objects and if crossmodal priming has similar/different effects in visual and haptic object recognition.

First of all, it is important to underline that this is one of the few studies in this research field that involves the use of novel 3D objects. In fact, in previous behavioural studies participants faced with 2D objects (Lederman et al., 1990; Prather et al., 2004), familiar objects (Craddock & Lawson, 2008; Easton et al., 1997a; Lacey & Campbell, 2006; Reales & Ballesteros, 1999) or very small objects which could be held in one hand (Newell et al., 2001; Lacey et al., 2007).

I asked participants to tell already explored objects from new ones, using different sensory modalities (vision or haptics) during the two phases (encoding and recognition). My hypothesis was that the sensory modalities order of presentation would have had different effect on recognition accuracy level. In particular, I expected visual priming to led to a higher recognition accuracy than the haptic one.

Concerning cross-modal recognition, I found that participants are accurate in distinguishing objects previously explored through another sensory modality from objects that were not explored before.

If I consider the mere accuracy, the recognition performance seems to be dependent from the sensory modality in which the comparison is made. In particular, visual objects encoding seems to lead to a higher haptic recognition level than the reverse (haptic encoding and visual recognition), suggesting a visual primacy in object representation. If I consider the sensitivity as index of recognition, that is a more accurate measurement with this kind of data because it takes into account the participants' tendency to answer "yes" or "no", although the result does not get to significance, it clearly expresses a trend in supporting a facilitating effect in haptic recognition when vision comes first. These results support those obtained by different authors (Streri & Molina, 1994; Lacey & Campbell, 2006).

Different suggestions may be proposed to explain these results. First of all, the way in which the two sensory modalities takes us to acquire information about the object could have a primary role. In fact, vision allows us to have an holistic/simultaneous perception of the object which could facilitate a subsequent sequential (haptic) recognition while, on the contrary, starting with a sequential encoding we have a series of "successive impressions" (Gibson, 1962) which could have a cost in visual recognition.

Another explanation comes from researches on visual and haptic categorization processes: they showed, in fact, that the two sensory modalities weight the object's features in different ways. Thus, when we look at the object we firstly concentrate on its shape, while when we touch the object other properties, like texture, capture our attention (Cooke et al., 2007; Gaißert et al., 2010; Gaißert & Wallraven, 2012). Also in this case, hence, exploring the object visually could have facilitated its shape recognition.

In the end, as suggested by several authors (De Volder et al., 2001; Zhang et al., 2004; Newman et al., 2005) visual imagery could play a leading role in haptic recognition.

It is necessary to underline that I didn't ask participants explicitly that they were going to face a recognition task, I just told them about this at the end of the encoding phase. So, it would be interesting to see if results change by expressly requesting participants to memorize items during encoding because they will be ask to recognize them later.

Concluding, according to results from my study and previous literature, vision seems to have a very important (probably primary) role in haptic recognition.

Considering that, it would be newsworthy to study if gaze patterns play some role in haptic object recognition (like for visual perception, imagination and recognition).

5.2 Experiment 2 – Unimodal (haptic) Recognition

The purpose of this experiment was to investigate how the visual system influence unimodal haptic object recognition. In particular, I was wondering if gaze patterns play some role in haptic object recognition.

My hypotheses are that haptic object recognition implies general visual processing and that, in particular, manipulating gaze patterns during encoding affects haptic object recognition.

5.2.1 Method

Participants: 72 healthy university students (44 men and 28 women) from Lawrence Technological University (Southfield, Michigan), with a mean age of 23.9 years (range 18-42 years), voluntarily participated in this experiment in exchange of university credit or a small amount of money. All of them read and signed the informed consent. Participants referred normal or corrected to normal sight and touch.

Materials and apparatus: Similar materials and apparatus like Experiment 1 were used. The above described 12 plaster structures were utilized also for this experiment. The table was covered by a black curtain to prevent participants from seeing the wooden plank and the items fixed on it. Moreover, the table was leaning against a wall and at participants' eyes height there was a wooden desk with a laptop on it. On the laptop screen, three out of four different experimental conditions were shown. For the fourth condition a blindfold was used.

Procedure: The experiment consisted of four between-subject conditions, the "Screen Color" condition, the "Fixed Cross" condition, the "Moving Cross" condition and the "Blindfolded" condition. Each participant took part in one of the conditions (18 participants – 11m and 7f - for each of the four conditions).

During the encoding phase, participants were asked to haptically explore six objects, tucking their hands under the curtain, for 20 seconds each, trying to memorize their shape for the successive recognition among twelve objects. They were allowed to freely manipulate the items with two hands and with no restrictions, except picking up the objects. After 20 seconds of exploration participants were asked to raise their

hands and to wait for the next object. This procedure was repeated for all the six sculptures. After haptic exploration was completed, participants left the room and waited for 2 minutes.

This process of encoding could take place in four different conditions. While exploring items with both hands, participants were asked to:

- ✓ Freely look at the laptop screen which changed color at different time intervals and to say "change" when the color changes → "Screen Color" condition
- ✓ Fix a cross on the center of the screen and to say "change" when its color changes → "Fixed Cross" condition
- ✓ Follow a cross on the screen with their eyes and to say "change" when its color changes → "Moving Cross" condition
- ✓ Keep the blindfold on → "Blindfolded" condition

The four conditions were randomized among participants.
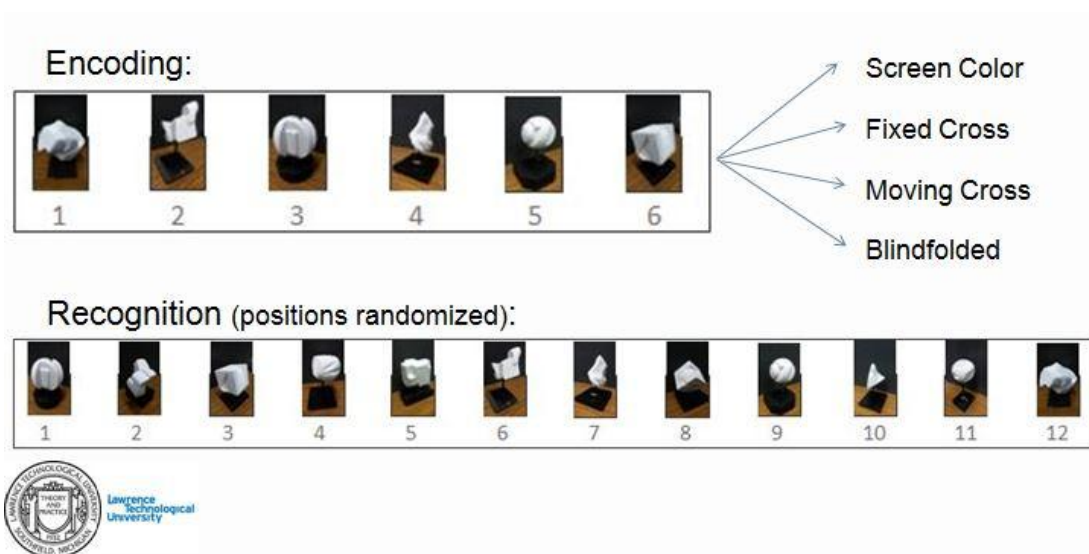
While the participant waited outside, all the 12 sculptures (the six that were just explored – target - plus six new ones – distracters – never presented before) were arranged on the wooden plank in a random order. The twelve sculptures order was randomized between participants. The participant was then led back into the room, and placed in front of the first sculpture's position on the table.

The recognition phase was the same for all participants. Participants haptically explored the first object under the curtain for the time they needed to provide an answer. They were asked to say "yes" if the object was one of the six touched before, "no" if they thought it was a new one. Once the answer was provided, they were led to the next item and the process repeated for each of the remaining 11 sculptures.

5.2.2 Analysis

At first I ran an ANOVA analysis with *Condition* (namely the experimental conditions: "Screen Color", "Fixed Cross", "Moving Cross" and "Blindfolded") as between-subjects factor and *Accuracy* as the dependent variable to test the different level of recognition accuracy for the four experimental conditions.

Then, signal detection theory (SDT) was applied to the accuracy data to calculate sensitivity (i.e. *d'*) and the presence of a response bias (i.e. criterion). *D-prime* and *Criterion* were calculated as reported in the previous experiment. To test the sensitivity and response bias of participants in the recognition task, I ran two ANOVA analyses with *Condition* as between-subjects factor and *d'* and *Criterion* as the dependent variables.

When necessary, Tukey HSD method was used for post-hoc analyses.

5.2.3 Results

Results show a significant effect for the factor "*Condition*", $F(3, 68) = 7.71$, $p < .001$, $\eta_p^2 = .254$. Post-hoc analyses indicated that accuracy for "Screen Color" condition (82.8%) was significantly higher than "Moving Cross" condition (66.2%; $p < .001$) and "Blindfolded" condition (72.2%; $p < .05$). Moreover, accuracy for "Fixed Cross" condition (78.2%) is significantly higher than accuracy for "Moving Cross" condition ($p < .01$).
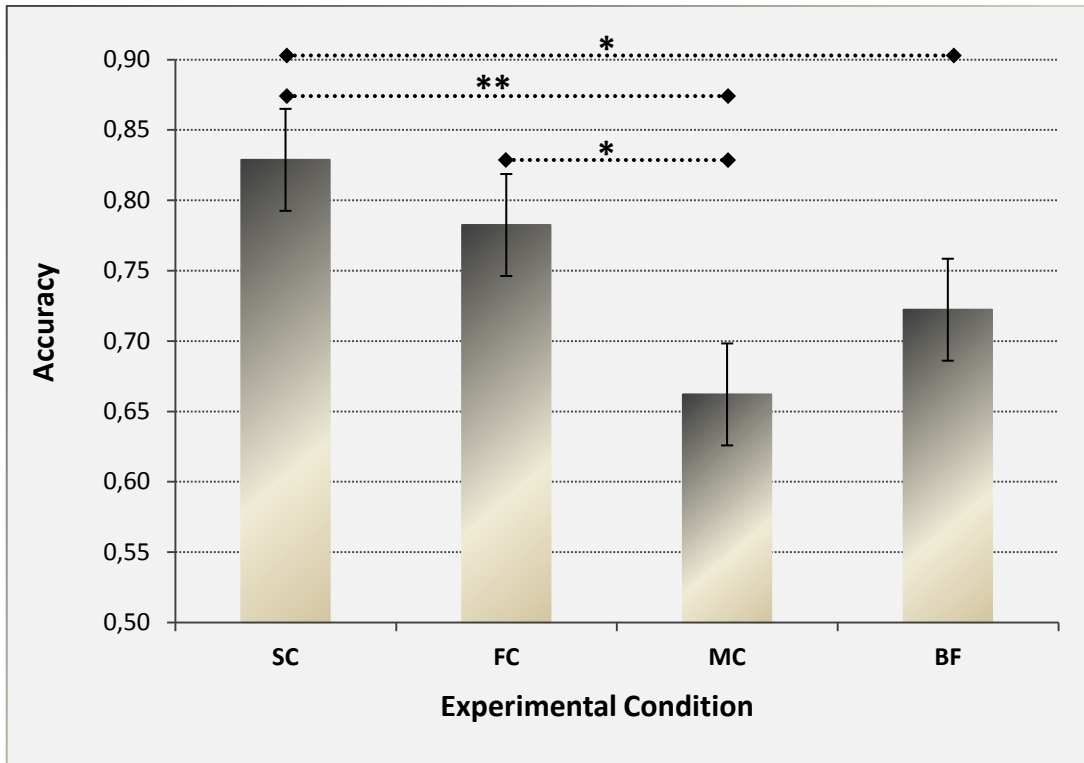
Note: *= p < .05; **= p < .001

Two single sample t-test were conducted to determine if d' and *Criterion* values were significantly different from zero. Results show that the d' values ($M = 1.99$, $SD = 1.93$) are significantly different from zero, indicating that participants perform above chance level [$t(71) = 8.71$, $p < .000$]. The *Criterion* values ($M = 0.11$, $SD = 0.74$) are not statistically different from zero, indicating that there is no response bias.

Concerning recognition sensitivity (SDT), results show a significant effect for *Condition, $F(3, 68) = 4.34$, $p = .007$.* Post-hoc analyses pointed that sensitivity for "Screen Color" condition (2.96) was significantly higher than "Moving Cross" condition (1.25; $p < .05$) and "Blindfolded" condition (1.25; $p < .05$) (See Table 1).

For criterion, results show no significant effects for *Modality*, $F(3, 68) = .16$, $p > .05$. This suggests that response approach (liberal vs. conservative) was not influenced by the experimental condition.

Table 1

Haptic Sensitivity for Different Experimental Conditions

		Experimental Conditions		
	SC	MC	FC	BF
	(2.96)	(1.25)	(2.38)	(1.25)
		p values		
SC		.022*	.754	.022*
MC	.022*		.214	1.00
FC	.754	.214		.215
BF	.022*	1.00	.215	

Note: *= p < .05

5.2.4 Discussion

The goal of this study was to understand how the visual system influence unimodal haptic object recognition. In particular, I was wondering if manipulating the natural gaze patterns during haptic object encoding would lead to different accuracy in recognition.

First of all, it is important to underline that this is the first study (in my knowledge) trying to investigate the role of gaze patterns in haptic perception, and it's one of the few studies in this research field that involves the use of novel 3D objects.

I asked participants to tell already explored objects from new ones using haptics both for encoding and recognition. My hypothesis was that haptic object encoding and recognition imply a visual processing and that for this process also gaze patterns are engaged.

Disrupting gaze patterns during haptic encoding resulted in a significant lower object recognition accuracy. In fact, when asked to follow a stimulus with their eyes during the encoding phase, participants were then less accurate in recognizing objects.

Analogous results were obtained when participants wore a blindfold during haptic object encoding.

For the first case, a possible explanation might be that disrupting free gaze patterns during object encoding could prevent subjects from forming a mental image of object in memory. Consideration about "mental imagery" when we deal with studies concerning recognition and memory are very important. Recognition, in fact, occurs when an object perception corresponds to a representation stored in memory (Bülthoff & Edelman, 1992). These results would be in line with several studies demonstrating that, in visual representation, we not only store the contents of a scene but also the gaze patterns we made to explore it (Laeng & Teodorescu, 2002; Johansson et al., 2006; Humphrey & Underwood, 2008). Eye movements during visual exploration obviously take part in scene encoding and we are facilitated in recovering it from memory by reproducing those movements. Could this be the case for haptic exploration also?

The most particular (unexpected) result for me is the difference between "Screen Color" and "Blindfolded" conditions, namely people perform worse when they are blindfolded than when they are asked to look at a laptop display and to say when its color changes during haptic object encoding. Perhaps because a frame of reference is missing? Or blindfolded participants don't have a strategy to follow? An answer comes from studies comparing non-informative vision (i.e., visual information that is not linked with the investigated one) and no vision (blindfolded) conditions.

In particular, Newport et al. (2002) supported that in circumstances favoring extrinsic (visual) coordinates, providing non-informative visual information significantly increases the accuracy of haptic perception. Providing non-informative vision may elicit the default use of an external reference frame to represent tactile stimuli. Moreover, spatial attention may be a component of non-informative vision as defined by the authors.

At this point, it would be interesting to pay attention to the kind of gaze patterns participants do for exploring the target items used for these experiments.

5.3 Experiment 3 – Eyetracking Data

The third experiment arises from the exploratory idea and with the general aim of investigating if there are differences in the gaze performances for the six target items used in the previous experiment. In particular, I wanted to understand if some items are better recognized than others and if gaze patterns are different for the best and the worst recognized stimuli.

5.3.1 Method

Participants: 12 healthy participants (7 men and 5 women) with a mean age of 28.2 years (range 23-32 years), volunteered to participate in this experiment. Informed consent was obtained from all participants. All of them reported normal or corrected to normal sight.

Materials and apparatus: Material was constituted by 18 images, obtained from three pictures for each of the six target items used in the previous two experiments (3x6), from three view angles (0°, 90°, 270°). These three angles were chosen because in the previous experiment participants were allowed to haptically explore items without any restriction, except picking them up. Gaze patterns (scanpaths) were registered with a SMI Eye Tracker RED500.

Procedure: The setup included a double monitor setup: one monitor for the operator and a second monitor to present the visual stimuli to participants. Participants entered
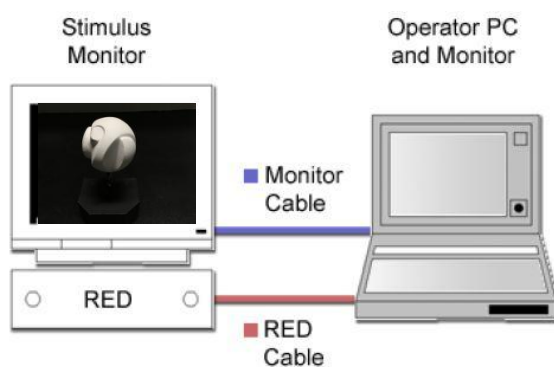
the experimental room and seated in front of the Stimulus Monitor in order to read instructions.

They were asked to visually explore the pictures of 6 abstract sculptures from three view angles and to try memorizing them. After making me sure that instructions were clear, we proceeded with the calibration, in order to adjust the software to the participant's eye characteristics. Participants looked at each sculpture's pictures for 10s (4s for 0°, 3s for 90° and 270°). At the end, I asked them

a brief description for each of the 6 sculptures they saw and remembered, just to verify they attentively looked at the items and to let them carry out a task.

5.3.2 Analysis

At first I ran an ANOVA analysis with *Item* (namely the 6 target items from the previous experiment) as within-subjects factor and *Accuracy* as the dependent variable to test the different level of recognition accuracy for the six target items.

Then, pairwise comparisons using paired t-tests (post-hoc analysis) were calculated to observe the differences between couples of items. Bonferroni was used as p value adjustment method.

Eye tracking data were calculated (number of fixations and saccades, fixation duration average, fixation dispersion average, saccade duration average and scanpath length) and compared for the 6 items. Then, heat map data views (attention map, in which different colors reveal the areas receiving more attention, being red the most, then yellow, green and blue) were created.

5.3.3 Results

Results show no significant effect for the factor "*Item*", $F(5,426) = 0.43$, $p > .05$. Actually, this analysis was run for the previous experiment, demonstrating that there is not any difference in item recognition, so all results can be related to the different experimental conditions.

However, I decided to proceed with the pairwise comparisons aiming to point out if some items were recognized better than others, albeit not significantly.

As we can see in Table 2, there is only one item (item 2A) that seems to be recognized worse than the others, while all the other items seem approximately equally recognized.

Table 2

Pairwise comparisons using paired t-tests

	ITEM 1A (0.67)	ITEM 2A (0.57)	ITEM 3A (0.76)	ITEM 4A (0.75)	ITEM 5A (0.75)	ITEM 6A (0.79)
			p-values			
ITEM 1A	-	1.00	1.00	1.00	1.00	1.00
ITEM 2A	1.00	-	0.33	0.28	0.34	0.13
ITEM 3A	1.00	0.33	-	1.00	1.00	1.00
ITEM 4A	1.00	0.28		-	1.00	1.00
ITEM 5A	1.00	0.34	1.00	1.00	-	1.00
ITEM 6A	1.00	0.13	1.00	1.00	1.00	-

As to eye tracking data, all dimensions (the number of fixations and saccades, the fixation duration average, the fixation dispersion average, the saccade duration average and the scanpath length) resulted to be almost the same for all of the 6 items.

The heat maps, instead, seem to show interesting suggestions (Figure 4). For almost all of the items, it seems that participants concentrate their attention on specific areas (red spots) when they study the objects trying to memorize them. For item 2A instead (and in a sense for item 6 too), an area not well identified seems to capture participants' attention.

5.3.4 Discussion

The goal of this study was to investigate if there are differences in the gaze performances for the six target items used in the previous experiment. In particular, I was wondering if gaze patterns are different for the best and the worst recognized stimuli.

As shown by the recognition accuracy for the six items, it seems that they are equally recognized by participants. This is a good point for the two previous experiments,

namely there are not items that could be memorized and recognized more easily than others, so results obtained can be addressed to participants' performances in the experimental conditions. Anyway, comparisons among couples of items indicated that item 2A is worse recognized than the others.

As to eye tracking data, the different dimensions calculated for the six items (number of fixations and saccades, fixation duration average, fixation dispersion average, saccade duration average and scanpath length) resulted to be almost the same. These results could be considered in line with haptic recognition, which did not show differences among items.

If we look at heat maps, however, we notice that when asked to memorize objects, participants concentrate their attention on specific areas/elements of the items. This is probably due to the fact that participants try to memorize peculiar elements for object recognition, besides the fact that they know they have very few seconds for storing information about the object. For item 2A this does not seem to be the case: it looks like participants try to explore all the object in search of a particular feature. Might we try to build a bridge between visual and haptic exploration? For a response, it would be useful to compare these eye tracking data with visual recognition data of items.
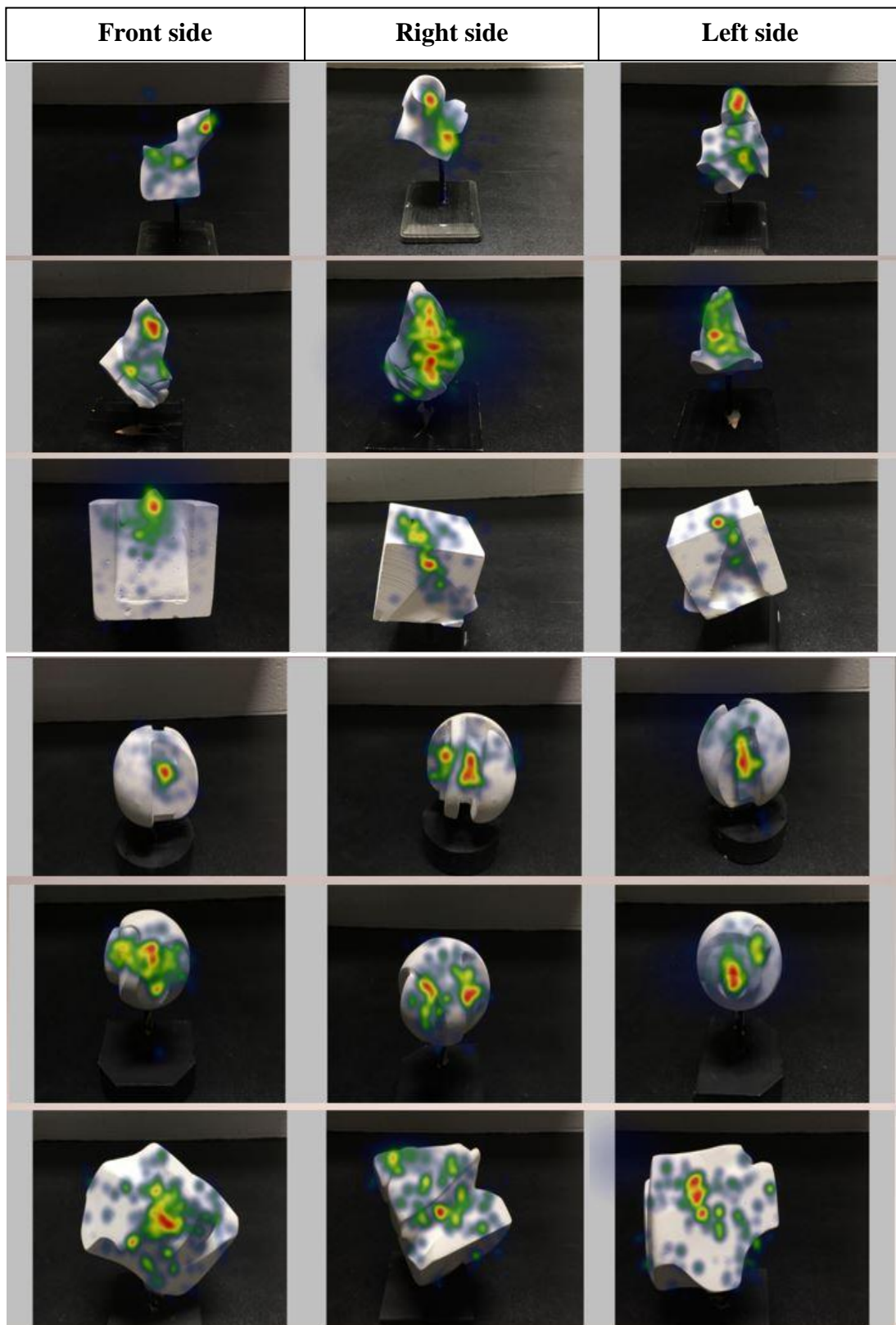
	Front side	Right side	Left side
Item 1A			
Item 2A			
Item 3A			
Item 4A			
Item 5A			
Item 6A			

Figure 4. Heat maps for the 6 target items

6. General discussions and Conclusions

The main goal of the present research was to inquire into the role of visual processing in haptic object perception and representation, using 3D abstract not labelable sculptures. Moreover, the idea was to try to propose an answer for the diatribe seeing opposing consideration about the existence of an amodal or supramodal object representation for visual and haptic systems. In fact, although it is true that several studies already supported the existence of a shared representation for vision and haptics, almost none of them demonstrated its being by using abstract 3D objects, which cannot provide any familiar prompt and, furthermore, all of them brought opposing supports for the representation mode.

To this end, I first analyzed the literature about how visual objects and haptic objects are represented and recognized, then I reviewed studies on crossmodal (visual-haptic and haptic-visual) objects perception and recognition. Relying on the cited literature, I carried out three experiments. As to the first one, the general purpose was to investigate if vision has a primary role in recognizing abstract 3D objects and if crossmodal priming has similar/different effects in visual and haptic object recognition. Based on available literature, the hypothesis was that visual priming would have led to a better haptic object recognition than the opposite condition (sensory modalities reversed). Results seem to support this hypothesis, claiming a cross-modal recognition asymmetry for vision and haptics, relying primarily on the visual modality, as already demonstrated by behavioral (Streri & Molina, 1994; Lacey & Campbell, 2006) and neuroimaging (Kassuba et al., 2013) studies. I have argued proposing different suggestions to explain these results. First of all, the way in which the two modalities acquire information about the object could have a primary role. In fact, vision allows us to have a holistic/simultaneous perception of the object which could facilitate a subsequent sequential (haptic) recognition, while starting with a sequential encoding could have a cost in simultaneous (visual) recognition. Another suggestion is the one supported by studies about categorization processes (Cooke et al., 2007; Gaißert et al., 2010) suggesting the idea that the

information about the shape is primarily important for the visual system but not for the haptic one which, instead, privileges the information about object texture (Cooke et al., 2007). In the end, I could suggest that visual imagery plays an important role in haptic recognition. According with several authors, indeed, visual imagery is fundamental in haptic object processing (Sathian et al., 1997; Zhang et al., 2004). This hypothesis is based on the observed LOC activation during visual imagery tasks with both sighted and blind people (De Volder et al., 2001; Newman et al., 2005). Considering results from this studies and previous literature, then, vision seems to have a very important (probably primary) role in haptic recognition.

With the second experiment (unimodal haptic recognition), I aimed to investigate how the visual system influence unimodal haptic object recognition. In particular, it would be newsworthy to see if gaze patterns play some role in haptic object recognition, like for visual perception, imagination and recognition. My hypothesis was that manipulating the natural gaze patterns during haptic object encoding would affect accuracy in recognition. In fact, results from this second experiment showed that participants were less accurate in recognizing objects when gaze patterns were disrupted. A suggested explanation might be that disrupting free gaze patterns during object encoding could prevent participants from forming a mental image of the object in memory. Consideration about "mental imagery" when we deal with studies concerning recognition and memory are very important. Recognition, in fact, occurs when an object perception corresponds to a representation stored in memory (Bülthoff & Edelman, 1992). These results seem to be in line with several studies demonstrating that, in visual representation, we not only store the contents of a scene but also the gaze patterns we made to explore it (Laeng & Teodorescu, 2002; Johansson et al., 2006; Humphrey & Underwood, 2008). Eye movements during visual exploration assist the scene encoding and reproducing those same eye movements at a later time can promote its remembrance. Could this be the case for haptic exploration also?

Analogous results were obtained when participants wore a blindfold during haptic object encoding. This could be a surprising result, but trying to imagine why, it might be suggested that a frame of reference is missing in this case or blindfolded participants don't have a strategy to follow. An answer comes from studies (Newport

et al., 2002) comparing non-informative vision with no-vision (blindfolded) conditions, which supports that providing non-informative visual information significantly increases the accuracy of haptic perception. Maybe because providing non-informative vision may elicit the default use of an external reference frame to represent tactile stimuli and, furthermore, spatial attention may be a component of non-informative vision (as defined by authors).

It is important to underline that this is the first study (to my knowledge) approaching the investigation of the role of gaze patterns in haptic perception, and that performance differences in different conditions can open new search paths about the mechanisms engaged in haptic perception and recognition.

At this point, it was interesting to observe the kind of gaze patterns participants do for exploring the target items used for these experiments, and this became the purpose of the third experiment. Moreover, I was wondering if gaze patterns are different for the best and the worst recognized stimuli. However, the six items seem to be equally recognized by participants and, in a sense, this could be reflected by eye tracking data, showing no differences among the calculated parameters (number of fixations and saccades, fixation duration average, fixation dispersion average, saccade duration average and scanpath length). Anyway, comparisons between item pairs indicated that as to haptic recognition item 2A is worse recognized than the others, so I tried to see if something analogous happens with eye tracking data. Heat maps showed that when asked to memorize objects, participants concentrate their attention on specific areas/elements of the items. For item 2A this does not seem to be the case: it looks like participants try to explore all the object in search of a particular feature. For a better exploration, it would be useful to compare these eye tracking data with visual recognition task data for items.

To conclude, taken together these results offer interesting insights into how we represent and recognize haptic object and how this is accompanied by concurrent visual processing mechanisms. In fact, although there is agreement on a shared representation for visuo-haptic object, it is not clear yet if the two sensory modalities are equally weighted and if the shared representation is amodal or supramodal.

Results obtained with this research seem to suggest an initially modal representation for haptic and visual systems, predominantly based on vision (for sighted individuals), then converging in a supramodal representation, where vision and haptics share object- information and influence each other. A possible mechanism for this, as already proposed by Riesenhuber and Poggio (1999) for visual object recognition, is the integration of various unimodal representations into a higher-order supramodal representation.

This is supported by several studies, like those showing a contraposition in object recognition between unimodal viewpoint-dependence from one side, and crossmodal viewpoint-independence from the other (Lacey et al., 2007; Ueda & Saiki, 2007). This result would suggest the existence of a different supramodal representation (Newell et al., 2005) involved in the processing of object shape information independently from the sensorial system encoding the information (Cattaneo et al., 2008). As to brain regions convoked by this possible higher-order representations, many authors indicated that the later occipital complex (LOC) seems to be activated by both visual and haptic stimuli basically in the same way (Amedi et al., 2001; Amedi et al., 2002; Prather et al. 2004).

Lacey et al. (2009a) collected all these results and other in the proposal of a model which describes visuo-haptic object representation. According to the authors, the multimodal representation is located in the LOC. This representation doesn't care about the input sensory modality and can be accessed by two different pathways (top-down and bottom-up) depending on whether the processed object is known or novel. The top-down pathway goes from prefrontal and parietal areas to the LOC and it is activated by familiar objects processing using object visual imagery, while the bottom-up pathway proceeds from the somatosensory cortex to the LOC and it is implicated by unfamiliar objects processing using spatial visual imagery (and the IPS; Lacey et al., 2009a).

Taken together, these results reflect the more and more shared idea of the existence of a supramodal brain instead of a unimodal inputs combination.

7. Future Directions

The results obtained from this research lead to some ideas and proposals for future studies.

As to crossmodal object recognition, it would be interesting to see what happen if we explicitly inform participants that they are going to face a recognition task.

As to the role played by gaze patterns in haptic recognition, it would be interesting to know:

- what kind of eye movements participants do during real object encoding and recognition
- if there are different gaze patterns for different sensory priming
- what results we would obtain by interfering with the recognition phase

In the end, for a complete view, it would be interesting to repeat these studies associating neuroimaging techniques.

Bibliography

Aleman A., van Lee L., Mantione M.H.M., Verkoijen I.G., De Haan E.H.F. (2001). Visual imagery without visual experience: Evidence from congenitally blind people. *Neuroreport,* 12: 2601–2604.

Alexander J.M., Johnson K.E., Schreiber B.E. (2002). Knowledge is not everything: Analysis of children's performance on a haptic comparison task. *Journal of Experimental Child Psychology*, 82(4): 341–366.

Amassian V.E., Cracco R.Q., Maccabee P.J., Cracco J.B., Rudell A.P., Eberle L. (1998). Transcranial magnetic stimulation in study of the visual pathway. *Journal of Clinical Neurophysiology*, 15: 288-304

Amedi A., Jacobson G., Hendler T., Malach R., Zohary E. (2002). Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cerebral Cortex*, 12: 1202–1212.

Amedi A., Malach R., Hendler T., Peled S., Zohary E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience*, 4: 324–330.

Amedi A., Raz N., Pianka P., Malach R., Zohary E. (2003). Early 'visual' cortex activation correlates with superior verbal memory performance in the blind. *Nature Neuroscience*, 6: 758–766.

Amedi A., von Kriegstein K., van Atteveldt N.M., Beauchamp M.S., Naumer M.J. (2005). Functional imaging of human crossmodal identification and object recognition. *Experimental Brain Research*, 166: 559–571.

Barsalou L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22: 577–600.

Bartolomeo P., Bachoud-Lévi A.C., De Gelder B., Denes G., Dalla Barba G. et al. (1998). *Neuropsychologia*, 36: 239-249

Behrmann M., Winocur G., Moscovitch M. (1992). Dissociation between mental imagery and object recognition in a brain-damaged patient. *Nature,* 359: 636-637.

Bértolo H., Paiva T., Pessoa L., Mestre T., Marques R., Santos R. (2003). Visual dream content, graphical representation and EEG alpha activity in congenitally blind subjects. *Cognitive Brain Research*, 15: 277-284.

Bértolo H. (2005). Visual imagery without visual perception? *Psicologica*, 26: 173:188

Biederman I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94: 115-147.

Biederman I. & Ju G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, 20: 38-64

Bliss I. & Hämäläinen H. (2005). Different working memory capacity in normal young adults for visual and tactile letter recognition task. *Scandinavian Journal of Psychology,* 46: 247–251.

Brandt S. & Stark L. (1997). Spontaneous Eye Movements During Visual Imagery Reflect the Content of the Visual Scene. *Journal of Cognitive Neuroscience*, 9(1): 27-38.

Bülthoff H.H., Edelman S., Sklar E. (1991). Mapping the generalization space in object recognition. *Investigative Ophthalmology and Visual Science*, 32, 996

Bülthoff H.H. & Edelman S. (1991). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences of the United States of America*, 89: 60-64

Bundesen C., Larsen A., Farrell J.E. (1981). Mental transformations of size and orientation. *Attention and Performance*, 9: 279-294.

Bushnell E.W. & Baxt C. (1999). Children's haptic and cross-modal recognition with familiar and unfamiliar objects. *Journal of Experimental Psychology: Human Perception and Performance,* 25: 1867–1881

Bushnell E.W. & Boudreau J.P. (1991). The development of haptic perception during infancy. In M. A. Heller & W. Schiff (Eds.), The psychology of touch (pp. 139–161). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Casey S.J. & Newell F.N. (2007). Are representations of unfamiliar faces independent of encoding modality? *Neuropsychologia,* 45: 506–513.

Cattaneo Z., Vecchi T., Cornoldi C., Mammarella I., Bonino D., Ricciardi E., Pietrini P. (2008). Imagery and spatial processes in blindness and visual impairment. *Neuroscience and Biobehavioral Review,* 32(8): 1346-60.

Chatterjee A. & Southwood M.H. (1995). Cortical blindness and visual imagery. *Neurology,* 45: 2189-2195.

Chen W., Kato T., Zhu X.H., Ogawa S., Tank D.W., Ugurbil K. (1998). Human primary visual cortex and lateral geniculate nucleus activation during visual imagery. *Neuroreport,* 9: 669–674.

Cohen L.G., Celnik P., Pascual-Leone A., Corwell B., Falz L. et al. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature*, 389: 180–183.

Cohen L.G., Weeks R.A., Sadato N., Celnik P., Ishii K., Hallett M. (1999). Period of susceptibility for cross-modal plasticity in the blind. *Annals of Neurology*, 45: 451–460.

Collignon O., Lassonde M., Lepore F., Bastien D., Veraart C. (2007). Functional cerebral reorganization for auditory spatial processing and auditory substitution of vision in early blind subjects. *Cerebral Cortex*, 17(2): 457–465.

Cooke T., Jäkel F., Wallraven C., Bülthoff, H. H. (2007). Multimodal similarity and categorization of novel, three-dimensional objects. *Neuropsychologia*, 45: 484–495.

Cooper E.E., Biederman I., Hummel J.E. (1992). Metric invariance in object recognition: a review and further evidence. *Canadian Journal of Psychology*, 46(2): 191-214

Craddock M. & Lawson R. (2008). Repetition priming and the haptic recognition of familiar and unfamiliar objects. *Perception & Psychophysics*, 70: 1350–1365.

Craver-Lemley C. & Reeves A. (1992). How visual imagery interferes with vision. *Psychological Review*, 99(4): 633–649.

D'Angiulli A. & Waraich P. (2002). Enhanced tactile encoding and memory recognition in congenital blindness. *International Journal of Rehabilitation Research*, 25(2): 143–145.

Damasio A.R., Tranel D., Damasio H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience*, 13: 89–109.

Deshpande G., X. Hu, Lacey S., Stilla R., Sathian K. (2010). Object familiarity modulates effective connectivity during haptic shape perception. *NeuroImage*, 49: 1991–2000.

Deshpande G., X. Hu, Stilla R., Sathian K. (2008). Effective connectivity during haptic perception: A study using Granger causality analysis of functional magnetic resonance imaging data. *NeuroImage*, 40: 1807–1814.

D'Esposito M., Detre J.A., Aguirre G.K., Stallcup M., Alsop D.C., Tippet L.J., Farah M.J. (1997). A functional MRI study of mental imagery generation. *Neuropsychologia*, 35: 725–730.

De Volder A.G., Toyama H., Kimura Y., Kiyosawa M., Nakano H. et al. (2001). Auditory triggered mental imagery of shape involves visual association areas in early blind humans. *NeuroImage*, 4(1): 129-139

Easton R.D., Greene A.J., Srinivas K. (1997a). Transfer between vision and haptics: Memory for 2-D patterns and 3-D objects. *Psychonomic Bulletin & Review*, 4: 403–410.

Easton R.D., Srinivas K., Greene A.J. (1997b). Do vision and haptics share common representations? Implicit and explicit memory within and between modalities. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 23: 153–163.

Engel S.A., Glover G.H., Wandell B.A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, 7: 181–192.

Farah M. (1989). The neural basis of mental imagery. *Trends in Neurosciences*, 12(10): 395-399

Feinberg T.E., Rothi L.J, Heilman K.M. (1986). Multimodal agnosia after unilateral left hemisphere lesion. *Neurology*, 36: 864–867

Ferreira F., Apel J., Henderson J.M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, 11: 405–410

Foulsham T. & Underwood G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, 8(2): 1-17

Gaißert N. & Wallraven C. (2012). Categorizing natural objects: a comparison of the visual and haptic modalities. *Experimental Brain Research*, 216(1): 123–134

Gaißert N., Wallraven C., Bülthoff H.H. (2008). Analyzing Perceptual Representations of Complex, Parametrically-Defined Shapes Using MDS. In: Ferre M. (eds) Haptics: Perception, Devices and Scenarios. EuroHaptics 2008. Lecture Notes in Computer Science, vol 5024. Springer, Berlin, Heidelberg

Gaißert N., Wallraven C., Bülthoff H.H. (2010). Visual and haptic perceptual spaces show high similarity in humans. *Journal of Vision*, 10(2)

Gallace A. & Spence C. (2009). The cognitive and neural correlates of tactile memory. *Psychological Bulletin*, 135(3): 380–406

Ganis G., Thompson W.L. & Kosslyn S.M. (2004). Brain areas underlying visual mental imagery and visual perception. *Cognitive Brain Research*, 20: 226–241

Gauthier I., Hayward W.G, Tarr M.J., Anderson A.W., Skudlarski P., Gore J.C. (2002). BOLD activity during mental rotation and view-dependent object recognition. *Neuron,* 34(1): 161–171

Gibbs R.W. (2006). Embodiment and cognitive science. New York: Cambridge University Press

Gibson J.J. (1962). Observations on active touch. *Psychological review*, 69(6): 477-491

Grill-Spector K., Kushnir T., Edelman S., Itzchak Y., Malach R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron*, 21: 191-202

Grill-Spector K., Kushnir T., Malach R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, 24(1): 187–203

Hatwell Y. (2003). Manual exploratory procedures in children and adults. In Y. Hatwell, A. Streri, & E. Gentaz (Eds.), Touching for knowing (pp. 67–83). Amsterdam: John Benjamins Publishing Co.

Haxby J.V., Grady C.L., Horwitz B., Ungerleider L.G., Mishkin M., et al. (1991). Dissociation of Spatial and Object Visual Processing Pathways in Human Extrastriate Cortex. *Proceedings of the National Academy of Sciences of the United States of America.* 88(5): 1621-1625

Haxby J.V., Grady C.L., Horwitz B., Salerno J., Ungerleider L.G. et al. (1993). Dissociation of Object and Spatial Visual Processing Pathways in Human Extrastriate Cortex. In Functional Organisation of Human Visual Cortex. Edited by Gulyàs B, Ottoson D, Roland P.E. Oxford: Pergamon Press; 329-340

Hebb D.O. (1968). Concerning imagery. *Psychological Review*, 75: 466–477

Heller M.A. (2000). Conclusions: The San Marino discussion. In M. A. Heller (Ed.), Touch, representation and blindness (pp. 183–217). New York: Oxford University Press.

Holšánová J., Hedberg B., Nilsson N. (1999). Visual and Verbal Focus Patterns when Describing Pictures. In: Becker W., Deubel H., Mergner T. (eds) Current Oculomotor Research. Springer, Boston, MA

Hubel D. & Wiesel T. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160 (1): 106–154.

Hummel J.E. & Biederman I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, 99(3): 480-517.

Humphrey G.K. & Khan S.C. (1992). Recognizing novel views of threedimensional objects. *Canadian Journal of Psychology*, 46: 170–90.

Humphrey K. & Underwood G. (2008). Fixation sequences in imagery and in recognition during the processing of pictures of real-world scenes. *Journal of Eye Movement Research,* 2(2): 1-15

James T.W, Humphrey G.K, Gati J.S, Menon R.S, Goodale M.A. (2002a). Differential effects of view on object-driven activation in dorsal and ventral streams. *Neuron*, 35: 793–801.

James T.W., Humphrey G.K., Gati J.S., Servos P., Menon R.S., Goodale M.A. (2002b). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia*, 40: 1706–1714.

James T.W., Culham J., Humphrey K., Milner D., Goodale M.A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain*, 126: 2463-2475

Johansson R., Holsanova J., Holmqvist K. (2006). Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science*, 30: 1053–1079.

Jolicoeur P. & Besner D. (1987). Additivity and interaction between size ratio and response category in the comparison of size-discrepant shapes. *Journal of Experimental Psychology: Human Perception & Performance*, 13: 478-487.

Kan I.P., Barsalou L.W., Solomon K.O., Minor J.K., Thompson-Schill S.L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology*, 20: 525–540.

Kanwisher N.G., Chun M.M., McDermott J., Ledden P.J. (1996). Functional imagining of human visual recognition. *Cognitive Brain Research*, 5: 55–67

Kaski D. (2002). Revision: is visual perception a requisite for visual imagery? *Perception*, 31: 717–731.

Kassuba T., Klinge C., Hölig C., Röder B., Siebner H.R. (2013). Vision holds a greater share in visuo-haptic object recognition than touch. *Neuroimage*, 65: 59–68.

Kiphart M.J., Hughes J.L., Simmons J.P., Cross H.A. (1992). Short term haptic memory for complex objects. *Bullettin of the Psychonomic Society*, 30: 212–214.

Klatzky R.L., Abramowicz A., Hamilton C., Lederman S. J. (2011). Irrelevant visual faces influence haptic identification of facial expressions of emotion. *Attention, Perception and Psychophysics*, 73(2): 521–530

Klatzky R.L., Lederman S.J., Metzger V. (1985). Identifying objects by touch: An "expert system". *Perception and Psychophysics*, 37: 299–302

Klatzky R.L. & Lederman S.J. (1993). Toward a computational model of constraint-driven exploration and haptic object identification. *Perception*, 22(5)

Knauff M., Kassubek J., Mulack T., Greenlee M.W. (2000). Cortical activation evoked by visual mental imagery as measured by fMRI. *Neuroreport*, 11: 3957–3962

Kohler S., Kapur S., Moscovitch M, Winocur G., Houle S. (1995). Dissociation of pathways for object and spatial vision: a PET study in humans. *NeuroReport,* 6: 1865-1868

Kosslyn, S.M. (1994). Image and Brain. MIT Press, Cambridge, MA.

Kosslyn S.M., Sukel K.E., Bly B.M. (1999a). Squinting with the mind's eye: effects of stimulus resolution on imaginal and perceptual comparisons. *Memory & Cognition*, 27(2): 276–287

Kosslyn S.M., Pascual-Leone A., Felician O., Camposano S., Keenan J.P., et al. (1999b). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science*, 284: 167–170

Kosslyn S.M., Thompson W.L., Kim I.J., Alpert N.M. (1995). Topographical representation of mental images in primary visual cortex. *Nature*, 378: 496–498

Kourtzi Z. & Kanwisher N.G. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, 20: 3310–3318

Kreiman G., Koch C., Fried I. (2000). Imagery neurons in the human brain. *Nature*, 408: 357-361

Lacey S. & Campbell C. (2006). Mental representation in visual/haptic crossmodal memory: evidence from interference effects. The *Quartely Journal of Experimental Psychology*, 59: 361–376

Lacey S., Flueckiger P., Stilla R., Lava M., Sathian K. (2010a). Object familiarity modulates the relationship between visual object imagery and haptic shape perception. *NeuroImage*, 49: 1977–1990

Lacey S., Hall J., Sathian K. (2010b). Are surface properties integrated into visuo-haptic object representations? *European Journal of Neuroscience*, 31: 1882–1888

Lacey S., Pappas M., Kreps A., Lee K., Sathian K. (2009b). Perceptual learning of view-independence in visuo-haptic object representations. *Experimental Brain Research*, 198: 329–337

Lacey S., Peters A., Sathian K. (2007). Cross-modal object representation is viewpoint-independent. *PLoS ONE*, 2(9): e890. doi: 10.1371/journal.pone0000890

Lacey S. & Sathian K. (2011). Multisensory object representation: insights from studies of vision and touch. *Progress in Brain Research,* 191: 165–176. doi: 10.1016/B978-0-444-53752-2.00006-0

Lacey S. & Sathian K. (2012). Representation of object form in vision and touch. In M. Wallace & M. Murray (Eds.). *The Neural Bases of Multisensory Processes*. Boca Raton, FL: CRC Press

Lacey S. & Sathian K. (2014). Visuo-haptic multisensory object recognition, categorization and representation. *Frontiers in Psychology*, 5, 730 doi: 10.3389/fpsyg.2014.00730

Lacey S., Stilla R., Sreenivasan K., Deshpande G., Sathian, K. (2014). Spatial imagery in haptic shape perception. *Neuropsychologia,* 60: 144-158 doi: 10.1016/j.neuropsycholopia.2014.05.008

Lacey S., Tal N., Amedi A., Sathian, K. (2009a). A putative model of multisensory object representation. *Brain Topography*, 21: 269–274. doi: 10.1007/s10548-009-0087-4

Laeng B., Bloem I.M., D'Ascenzo S., Tommasi L. (2014). Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition,* 131(2): 263–283

Laeng B. & Teodorescu D.S. (2002). Eye scanpaths during visual imagery re-enact those of perception of the same visual scene. *Cognitive Science*, 26: 207-231

Lashley K. (1948). The Mechanism of Vision, XVIII, Effects of Destroying the Visual "Associative Areas" of the Monkey, *Journal Press*, Provincetown MA, pp. 302-06

Lederman S.J. & Klatzky R.L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19: 342–368

Lederman S.J. & Klatzky R.L. (1990). Haptic classification of common objects: Knowledge-driven exploration. *Cognitive Psychology*, 22: 421-459

Lederman S.J. & Klatzky R.L. (1993). Extracting object properties through haptic exploration. *Acta Psychologica,* 84: 29-40

Lederman S.J., Klatzky R.L., Chataway C., Summers C. (1990). Visual mediation and the haptic recognition of two-dimensional pictures of common objects. *Perception & Psychophysics*, 47(1): 54–64

Loomis J.M. & Lederman S.J. (1986). 'Tactual perception'. In: K. Boff, L. Kaufman and J. Thomas (Eds.), Handbook of perception and human performance (pp. 31-41). Wiley: New York

Lowe D.G. (1985). The SCERPO vision system. *Perceptual Organization and Visual Recognition*, pp133-148

Lucan J.N., Foxe J.J., Gomez-Ramirez M., Sathian K., Molholm S. (2011). Tactile shape discrimination recruits human lateral occipital complex during early perceptual processing. *Human Brain Mapping*, 31: 1813–1821

Malach R., Reppas J.B., Benson R.R., Kwong K.K., Jiang H. et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 92: 8135–8139

Mazer J.A. & Gallant J.L. (2000). Object recognition: seeing us seeing shapes. *Current Biology*, 10: 668–670

Millar S. (1997). Theory, experiment and practical application in research on visual impairment. *European Journal of Psychology of Education*, 12(4): 415–430.

Mishkin M., Ungerleider L.G., Macko K.A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, 6: 414-417

Morrongiello B.A., Humphrey G.H., Timney B., Choi J., Rocca P.T. (1994). Tactual object exploration and recognition in blind and sighted children. *Perception*, 23(7): 833–848

Moscovitch M., Winocur G., Behrmann M. (1999). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, 9: 555–604

Newcombe F., Ratcliff G., Damasio H. (1987). Dissociable visual and spatial impairments following right posterior cerebral lesions: Clinical, neuropsychological and anatomical evidence. *Neuropsychologia*, 25: 149-61

Newell F.N., Ernst M.O., Tjan B.S., Bülthoff H.H. (2001). View dependence in visual and haptic object recognition. *Psychological Science*, 12: 37–42

Newell F.N. & Findlay J.M. (1997). The effect of depth rotation on object identification. *Perception,* 26: 1231–1257.

Newell F.N., Woods A.T., Mernagh M., Bülthoff H.H. (2005). Visual, haptic and crossmodal recognition of scenes. *Experimental Brain Research*, 161: 233–242.

Newman S.D., Klatzky R.L., Lederman S.J., Just M.A. (2005). Imagining material versus geometric properties of objects: An fMRI study. *Cognitive Brain Research*, 23: 235–246.

Newport R., Rabb B., Jackson S.R. (2002). Noninformative vision improves haptic spatial perception. *Current Biology*, 12(19): 1661-1664

Noordzij M.L., Zuidhoek S., Postma A. (2007). The influence of visual experience on visual and spatial imagery. *Perception*, 36: 101–112

Noton D. & Stark L. (1971a). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, 11: 929–942

Noton D. & Stark L. (1971b). Scanpaths in eye movements during perception. *Science*, 171: 308–311

Palmer S., Rosch E., Chase P. (1981). Canonical perspective and the perception of objects. In: Long J.B., Baddeley A.D., editors. In Attention and Performance IX. Hillsdale, NJ: *Lawrence Earlbaum Associates*; pp. 135–151

Pascual-Leone A., Amedi A., Fregni F., Merabet L.B. (2005). The plastic human brain. *Annual Review of Neuroscience*, 28: 377–401

Peissig J.J. & Tarr M.J. (2007). Visual object recognition: Do we know more now than we did 20 years ago? *Annual Review of Psychology*, 58: 75–96

Peltier S., Stilla R., Mariola E., LaConte S., X. Hu, Sathian K. (2007). Activity and effective connectivity of parietal and occipital cortical regions during haptic shape perception. *Neuropsychologia,* 45: 476–483

Pensky A.E.C., Johnson K.A., Haag S., Homa D. (2008). Delayed memory for visual-haptic exploration of objects. *Psychonomic Bullettin & Review*, 15(3): 574–580

Pietrini P., Furey M.L., Ricciardi E., Gobbini M.I., Wu W.H., Cohen L., Guazzelli M., Haxby J.V. (2004). Beyond sensory images: object-based representation in the human ventral pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 101: 5658–5663

Prather S.C., Votaw J.R., Sathian K. (2004). Task-specific recruitment of dorsal and ventral visual areas during tactile perception. *Neuropsychologia*, 42: 1079–1087

Reales J.M. & Ballesteros S. (1999). Implicit and explicit memory for visual and haptic objects: cross-modal priming depends on structural descriptions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(3), 644–663

Riesenhuber M. & Poggio T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2: 1019–1025

Sadato N. (2005). How the blind "see" Braille: lessons from functional magnetic resonance imaging. *Neuroscientist*, 11: 577–582

Sadato N., Okada T., Honda M., Yonekura Y. (2002). Critical period for cross-modal plasticity in blind humans: a functional MRI study. *Neuroimage*, 16: 389–400

Sadato N., Pascual-Leone A., Grafman J., Ibanez V., Deiber M.P., Dold G., Hallett M. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, 380: 526–528

Sathian K. (2004). Modality, quo vadis?. *Behavioral and Brain Sciences*, 27(3): 413–414

Sathian K. (2005). Visual cortical activity during tactile perception in the sighted and the visually deprived. *Developmental Psychobiology*, 46: 279–286.

Sathian K. & Lacey S. (2007). Journeying beyond classical somatosensory cortex. *Canadian Journal of Experimental Psychology*, 61: 254–264

Sathian K. & Stilla R. (2010). Cross-modal plasticity of tactile perception in blindness. *Restorative Neurology and Neuroscience*, 28: 271–281

Sathian K., Zangaladze A., Hoffman J.M., Grafton S.T. (1997). Feeling with the mind's eye. *Neuroreport*, 8: 3877–3881

Sato, T. (1989). Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake monkeys. *Experimental Brain Research*, 77: 23–30

Sereno M.I., Dale A.M., Reppas J.B., Kwong K.K., Belliveau J.W. et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268: 889-893

Simpkins K.E. (1979). Tactual discrimination of shapes. *Journal of Visual Impairment & Blindness*, 73: 93–101

Stilla R., Hanna R., X. Hu, Mariola E., Deshpande G., Sathian K. (2008). Neural processing underlying tactile microspatial discrimination in the blind: A functional magnetic resonance imaging study. *Journal of Vision*, 8: 1–19

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

Stoesz M., Zhang M., Weisser V.D., Prather S.C., Mao H., Sathian K. (2003). Neural networks active during tactile form perception: Common and differential activity

during macrospatial and microspatial tasks. *International Journal of Psychophysiology*, 50: 41–49

Streri A. & 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

Tanaka K. (1993). Neuronal mechanisms of object recognition. *Science*, 262(5134): 685-688

Tanaka K. (1996). lnferotemporal cortex and object vision. *Annual Review of Neuroscience*, 19: 109-l 39

Tarr, M.J. & Pinker, S. (1989). Mental rotation and orientation dependence in shape recognition. *Cognitive Psychololy*, 21(2): 233–282

Thomas L. & Lleras A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin and Review*, 14: 663–668

Tootell R.B.H., Dale A.M., Sereno M.I., Malach R. (1996). New images from human visual cortex. *Trends in Neurosciences*, 19(11): 481 -489

Tootell R.B.H., Hadjikhani N., Mendola J.D., Marett S., Dale A.M. (1998). From retinotopy to recognition: fMRI in human visual cortex. *Trends in Cognitive Sciences*, 2: 174–183

Uhl F., Kretschmer T., Lindinger G., Goldenberg G., Lang W., Oder W., Deecke L. (1994). Tactile mental imagery in sighted persons and in patients suffering from peripheral blindness early in life. *Electroencephalography and Clinical Neurophysiology*, 91: 249-255

Ueda Y. & Saiki J. (2007). View independence in visual and haptic object recognition. *Japanese Journal of Psychonomic Science*, 26: 11–19

Ueda Y. & 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

Ullman S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition*, 32: 193-254

Ungerleider L.G. & Haxby J.V. (1994). 'What' and 'where' in the human brain. *Current Opinion in Neurobiology*, 4(2), 157–165

Ungerleider L.G. & Mishkin M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield, *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press

Wang G., Tanifuji M., Tanaka K. (1998). Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. *Neuroscience Research*, 32: 33–46

Whitagen A., Kappers A.M.L., Vervloed M.P.J., Knoors H., Verhoeven L. (2012). Haptic object matching by blind and sighted adults and children. *Acta Psychologica*, 139: 261-271

Wojciulik E. & Kanwisher N. (2000). Visual attention: insights from brain imaging. *Nature Reviews Neuroscience*, 1: 91–100

Woods A.T., Moore A., Newell F.N. (2008). Canonical views in haptic object perception. *Perception*, 37: 1867–1878

Woods A.T., O'Modhrain S., Newell F.N. (2004). The effect of temporal delay and spatial differences on crossmodal object recognition. *Cognitive, Affective & Behavioral Neuroscience*, 4(2): 260–269.

Yildirim I. & Jacobs R.A. (2013). Transfer of object category knowledge across visual and haptic modalities: experimental and computational studies. *Cognition,* 126: 135–148

Zhang M., Weisser V.D., Stilla R., Prather S.C., Sathian K. (2004). Multisensory cortical processing of object shape and its relation to mental imagery. *Cognitive, Affective & Behavioral Neuroscience*, 4: 251–259